Lappeenranta University of Technology

School of Engineering Science
Machine Vision and Pattern Recognition Laboratory

CTU CZECH TECHNICAL UNIVERSITY IN PRAGUE

Faculty of Electrical Engineering
Department of Cybernetics

Blood vessel segmentation in the analysis of retinal and diaphragm images

Doctoral Thesis

written under a joint supervision (cotutelle) agreement between Lappeenranta University of Technology, Finland and Czech Technical University in Prague, the Czech Republic.

Ing. Pavel Vostatek

Prague, November 2017

PhD. Study Programme: P2612 - Electrical Engineering and Information Technology
Branch of Study: 3902V035 - Artificial Intelligence and Biocybernetics

Czech Technical University supervisor: **Doc. Daniel Novák, Ph.D.**
Lappeenranta University of Technology supervisor: **Prof. Lasse Lensu, Ph.D.**

Preface

Many people have helped and supported me during the course of the studies and work on this thesis. The work was divided between two universities in two countries which gave me the opportunity to meet many people who deserve my gratitude.

First of all, I would like to thank my supervisors throughout the years – Lasse Lensu, Daniel Novák and Olga Štěpánková – who provided kind and very supportive leading while offering their broad knowledge of the field. Their support was very valuable.

Lenka Hellebrandová allowed be to proceed further with the diaphragm work by asking me to process data in her study. My coauthors Hannu Uusitalo and Ela Claridge provided valuable comments on my papers regarding the retinal image processing. My gratitude goes to all these people.

I would like to thank to my coworkers and friends at the both MVPR laboratory at LUT and NIT laboratory at CTU for creating nice and enjoyable work atmosphere. Those people are Eduard Bakštein, Tomáš Sieger, Jirka Wild, Jirka Anýž, Jakub Schneider, Aidin Hassanzadeh, Sahar Zafari, Lauri Laaksonen and Toni Kuronen among others. Many thanks to Heikki Kälviäinen who was the first person from LUT I met and forwarded my request for studying at LUT further.

There are many people who made my stay in Finland very special and left many great memories. I would like to thank people from my Erasmus group and Daniela Zamora of the first semester for many special moments and travels. Further thanks go to Nicolas Taba, Camille Bajeux, Margaux Mauduit, Tommi Johansson, Yongyi Wu, Victoria Palacin Silva, Michal Genserek and many many others for the whole atmosphere during my stay in Finland.

Finally, many thanks to my family and friends in Prague for the most valuable things in life.

Prague, November 2017

Pavel Vostatek

Abstract

The segmentation and characterization of structures in medical images represents an important part of the diagnostic and research procedures in medicine. This thesis focuses on the characterization methods in two application fields that make use of two imaging modalities. The first topic is the characterization of the blood vessel structure in the human retina and the second is the characterization of diaphragm movement during breathing. The imaged blood vessel structures are considered important landmarks in both applications.

The framework for the retinal image processing and analysis starts with the testing of five publicly available blood vessel segmentation methods for retinal images. The parameters of the methods are optimized on five databases with the ground truth for blood vessels. An approach for predicting the method parameters is proposed based on the optimization results. The parameter prediction approach is then applied to obtain vessel segmentation on a new database and an automatic approach to the blood vessel classification and computation of the arteriovenous ratio is proposed and evaluated on the new database.

The framework for the diaphragm image processing and analysis is based on the measurement of diaphragm motion. The motion is characterized by a set of features quantifying the amplitude and frequency of the breathing pattern, as well as a portion of the non-harmonic movements that occur. In addition, a set of static features like the diaphragm slope and height are proposed. Two approaches for the motion measurement are proposed and compared. A statistical evaluation of the proposed features is performed by comparing measurements from people with and without spinal findings.

The results from the retinal image processing and analysis revealed the possibility of the successful prediction of the parameters of the blood vessel segmentation methods. The automatic approach for the automatic arteriovenous ratio estimation revealed a stronger association with blood pressure than the manually estimated ratio. The results from the diaphragm image processing and analysis confirmed differences in the position, shape and breathing patterns between the healthy people and people suffering from spinal findings. The blood vessel structure was shown to be a reliable marker for characterizing the diaphragm motion.

Keywords: image processing, segmentation, classification, retina, blood vessels, arteries, veins, arterio-venous ratio, diaphragm, breathing, spinal findings, low back pain, motion estimation

Abstrakt

Segmentace a charakterizace struktur v lékařských obrazových datech jsou důležitou součástí diagnostických a výzkumných procedur v medicíně. Tato dizertace se zabývá charakterizací struktur ve dvou oblastech, kde každá využívá jiný způsob snímání obrazu. První oblastí je charakterizace cévní struktury viditelné v sítnici oka. Druhou oblastí je charakterizace pohybu bránice během dechového cyklu. Cévní struktura viditelná jak v sítnici, tak v bránici, slouží v obou případech jako významná pomocná struktura.

V části zabývající se zpracováním snímků sítnice je nejdříve testováno pět metod, s veřejně dostupnou implementací, které byly navrženy pro segmentaci v sítnici viditelných cév. Parametry každé metody jsou optimalizovány na pěti veřejně dostupných databázích snímků sítnice, které jsou k dipozici s referenční segmentací cév. Výsledky optimalizace jsou pak využity pro návrh predikčního algoritmu pro odhad parametrů segmentace snímků v libovolné databázi. Predikční algoritmus je následně využit k segmentaci cév v nové databázi a segmentované cévy jsou využity pro návrh a validaci nového systému pro výpočet poměru šířky žil a tepen.

V části zabývající se zpracováním obrazů bránice je měřen a charakterizován pohyb bránice. U pohybu je měřen zdvih bránice a frekvence dechu. Také je umožněno určení poměru neharmonického pohybu, který bránice vykoná. Dále jsou měřeny statické příznaku, jako jsou sklon a výška bránice v hrudníku. Dva přístupy k měření samotného pohybu bránice jsou navrženy a porovnány. Je provedeno statistické porovnání příznaků mezi skupinami lidí bez nálezů na páteři a s nálezy.

Výsledky analýzy snímků sítnice prokázaly možnost predikce parametrů u metod pro segmentaci cév v sítnici i s aplikací navrženého predikčního systému. Navržený systém pro výpočet poměru šířky žil a tepen prokázal lepší asociaci mezi automaticky vypočtenou hodnotou tohoto poměru a krevním tlakem, než ručně počítané hodnoty. Výsledky analýzy obrazů bránice potvrdily rozdíly v poloze, tvaru a dechových vzorcích mezi zdravými jedinci a jedinci s nálezy na páteři. Cévy viditelné v bránici byly shledány jako spolehlivá záchytná struktura k měření pohybu bránice.

Acc	accuracy
AI	artificial intelligence
AMD	age-related macular degeneration
AP	anteroposterior
AUC	area under the ROC curve
AV	arteriovenous
AVR	arteriovenous ratio
ARIA	ARIA database
B	blue
BMI	body mass index
CHASEDB1	CHASEDB1 database
CLAHE	contrast limited adaptive histogram equalization
COPD	chronic obstructive pulmonary disease
COSFIRE	'combination of a shifted filter responses'
CRAE	central retinal artery equivalent
CRVE	central retinal vein equivalent
CV	computer vision
DIARETDB1	DIARETDB1 diabetic retinopathy database
DoG	difference of Gaussians
DR	diabetic retinopathy
DRIVE	Digital Retinal Images for Vessel Extraction
EM	expectation maximization
FFT	fast Fourier transform
FNR	false-negative rate
FOV	field of view
FPR	false-positive rate
G	green
GMM	Gaussian mixture model
GT	ground truth
GUI	graphical user interface
H	hue
HRF	HRF database
HSV	hue, saturation, value

IUWT	isotropic undecimated wavelet transform
k-NN	K-nearest neighbours
LBP	low back pain
LDA	linear discriminant analysis
LMSE	least mean square error
LRV	Likelihood Ratio Vesselness
LS-SVM	Least Squares Support Vector Machines
MCC	Matthew's correlation coefficient
MRI	magnetic resonance imaging
NN	neural network
OD	optic disc
ODD	optic disc diameter
PCA	principal component analysis
PDF	probability density function
PPV	positive predictive value
PR	pattern recognition
QDA	quadratic discriminant analysis
R	red
RGB	red, green and blue
RMSE	root mean square error
ROC	receiver-operating characteristic
ROI	region of interest
ROM	range of motion
ROP	retinopathy of prematurity
S	saturation
SD	standard deviation
Sn	sensitivity
Sp	specificity
STARE	Structured Analysis of the Retina
SVM	support vector machines
TNR	true-negative rate
TPR	true-positive rate
V	value
VAS	visual analog scale
VC	vital capacity

$\{\}$	a set
$max(.,.)$	a maximum of the inputs
$min(.,.)$	a maximum of the inputs
$[n_1, n_2]$	a closed interval between n_1 and n_2
$[n_1..n_2]$	a closed interval of integers between n_1 and n_2

Chapter I

Introduction

The continuous increase in the workload capacity of computer systems, the ease of sharing knowledge and data, and miniaturization are paving the way for computerized automated systems to solve increasingly complex tasks with an increasing amount of autonomy. In medicine, many use cases can be found for automated systems, beginning with blind people being able to use smart phones, cochlear implants allowing deaf people to hear and ending with systems like IBM's Watson being applied in reasoning about clinical diagnosis. Together with the growing computation power, the growing amount of data being recorded, stored and shared allows for the automated systems to surpass the deduction possibilities of humans in an increasing number of applications [1, 2]. In the medical field this allows for increasing proficiency of the computerized systems in disease diagnosis, prediction or disease prevention [3].

However, modern society, with its diversion from a traditional human lifestyle, suffers from the higher prevalence of a number of diseases. For example, related to the context of this thesis are diabetes and other diseases that manifest themselves in the eyes, which are among the leading cause of blindness today [4]. Another example is the raising prevalence of low back pain (LBP) [5]. These cases are considered important factors behind the motivation for the understanding and development of methods of early diagnosis, as well as methods that would allow decreasing the burden brought about by the diseases.

Computer vision (CV) and pattern recognition (PR) play important role together among the methods used within the automated systems and this thesis describes the employment of PR methods – medical image processing in particular – with a focus on the segmentation and diagnosis of vessel-like structures in two application areas. The first is in retinal image analysis where the CV and PR methods can and are expected to provide objective assessment of the retina [4]. The second is analysis of the diaphragm (the respiratory muscle) images where the CV and PR methods can help to provide a framework for understanding the principles behind the role of the diaphragm in establishing body posture, and its role in LBP and the prevention of the pain. The following paragraphs map out the general background of the clinical significance of retinal image processing and about LBP (with its connections to the diaphragm).

Visual examination of the retina provides a non-invasive view into the eye and at the same time into the central nervous system [4]. One of the distinct features of the retina are the blood vessels, whose structure is an important indicator of diseases such as diabetes, hypertension and cardiovascular disease [6]. Retinal imaging has been used to characterize the vessel structure, and diagnose, monitor and document abnormal conditions [7]. With the current technology, it is already possible to produce quantitative information of signs of eye diseases like diabetic retinopathy and glaucoma, as well as many cardiovascular and neurovascular diseases. A review of retinal imaging and its medical implications has been provided in [4].

To diagnose incipient abnormalities and diseases in their early stages, screening programmes with systematic protocols are being implemented for groups at risk. As the screening programmes become more extensive, the amount of data increases and, in many cases, manual diagnosis becomes a bottleneck. To remedy the problem with the increasing workload, computer-aided diagnosis tools can be used to provide access to retinal images and enable high-throughput workflows for the screening programs [4]. To enable automatic or semi-automatic image analysis and the structural characterization of the blood vessels, various approaches have been proposed for segmenting the vessels from retinal images: see [8] and [9]. A review of general vessel extraction techniques has been published in [10].

LBP is a very frequent phenomenon, while at the same time it is not easy to explain its origin. LBP has a very generic basis and numerous studies have tried to identify the most common source of the pain in the lower back, but no clear connection between LBP and the commonly believed sources like occupational poses or obesity, was concluded [11]. By contrast, evidence exits that the presence of respiratory disease is a stronger predictor for LBP than other established factors [11]. The diaphragm, besides its respiratory function, also has a role in body stabilization and altered diaphragm motion patterns in patients suffering from LBP [12, 13] are a validated fact. However, the exact mechanisms behind the role of the diaphragm in the genesis or suppression of LBP remains unclear. Characterization of the motion of the diaphragm in a objective way and the identification of its respiratory and postural component should improve our understanding of the field.

The usage of signal processing methods for separating the diaphragm motion into postural and respiratory parts is a novel method in the characterization of the diaphragm's kinematics and it could be a helpful step toward automatic characterization of the diaphragm. Our work should help in understanding the role of the diaphragm in stabilizing mechanisms and the possibilities for explaining the altered diaphragm movements [14, 15]. Because state-of-the-art methods usually only involve measurement of the static diaphragm, the dynamic properties are not well understood. Also, the visual inspection of the breathing patterns is an ideal target for the signal processing algorithms that will raise the precision of diaphragm assessment.

1.1 Objectives

The segmentation of vessel-like structures is connecting all topics in the thesis. First, a review of the available methods for the blood vessel segmentation in the retina is done and their performance is compared. Then, these methods are used for the characterization of

the retinal topology and to support motion tracking in the processing of the diaphragm images.

In the area of retinal image analysis, the thesis focuses on the blood vessel segmentation and the characterization of the blood vessel structure in color fundus photographs. Because the blood vessel segmentation in retinal images is a topic of active research [8], instead of proposing a new method for the segmentation task, it was decided to review the existing segmentation methods for which there is an implementation available and perform a controlled comparison of them. With the blood vessel segmentation methods in hand, this work aims to propose a framework for automatic estimation of arteriovenous ratio (AVR) – a commonly used measure for the characterization of the retina. The objectives of the first part of the thesis are as follows:

1. To review the existing retinal blood vessel segmentation methods for which there is an implementation available. In addition, to prepare an experiment to optimize the method parameters on various retinal databases with ground truth (GT) for the blood vessels and compare the method performance to each other and to the performance of the state-of-the-art methods.

2. To propose a model for setting the methods' parameters when applied to new retinal image databases. The results from the preceding optimization experiment are used for this.

3. To propose a framework for the automatic estimation of AVR and validate it on a new database where relevant clinical measurements from the subjects are available. The parameter prediction model is used to obtain the vessel segmentation on the database. The validation of the system is performed by comparison between associations with subjects' blood pressure and the AVR measured manually and by the proposed system.

The analysis of the diaphragm is aimed at the characterization of the diaphragm movement and shape to allow the investigation of differences between healthy people and people suffering from LBP or having spine findings. Dynamic series of magnetic resonance imaging (MRI) images of the diaphragm while breathing are used as the input to the processing. The particular objectives of the second part of the thesis are as follows:

4. To propose a set of features to characterize the motion and shape of the diaphragm in the body. The features allow for the investigation of the diaphragm motion in patients suffering from LBP and their comparison to healthy subjects. Simple measurement of the diaphragm's motion, defined by the area delineated by the diaphragm contour while moving, is used. The diaphragm contour is delineated manually.

5. To investigate and validate the usability of the vessel structure visible in the diaphragm for the motion assessment. The methods gathered for the retinal blood vessel segmentation are used to segment the vessels in the diaphragm. Tracking of the vessel structure will then be used to obtain accurate measurement of the diaphragm's motion.

1.2 Contributions

Contributions to the state of the art are as follows:

- A controlled comparison of the selected blood vessel segmentation methods while their parameters have been optimized was performed.
- The parameter prediction of the blood vessel segmentation methods is a novel approach to the best of our knowledge.
- The parts concerning the optimization of the parameters and the proposal of the predictive models were published as a conference paper [16] and as a journal paper [17].
- A framework for the classification of the segmented vessels into arteries and veins was proposed, building on the knowledge from the other state-of-the-art methods. A novel method for selecting a subset of the classified arteries and veins, and their combination was established in order to compute the AVR – a widely used measure for retinal vessel quantification – and validated by an assessment of associations between the AVR and the clinical data.
- The contributions of the diaphragm processing include an assessment of the non-respiratory diaphragm function resulting from the isolation of the diaphragm motions that are not related to respiration. This approach was used to process a set of data gathered from patients suffering from LBP and led to the proposal of various features that can be used to distinguish between the diaphragm movement of healthy subjects and that of patients suffering from LBP.
- Processing of the diaphragm image sequences was published as a conference paper [18] and as a journal paper [19], and the methods have also been applied in [20].
- Data from [20] were further used to improve the method in order to automatically obtain the diaphragm motion by registration of the segmented blood vessels in the diaphragm.

A detailed list list of the contributions of the thesis and the author's list of publications is presented in Chapter 6.

1.3 Thesis outline

Chapter 2, 'Medical and technical background', gives background information on the image processing applied in both modalities that are studied in this work – retinal and diaphragm images. The anatomical background of both fields is provided.

Chapter 3, 'Retinal blood vessel segmentation methods', gives an overview of the available methods for the vessel segmentation in colour retinal fundus photographs, which are available in the form of an implementation. Then the available retinal databases which contain ground truth information for the vessels are reviewed. Setup of the parameter

optimization experiment is defined and the results of the parameter optimization are presented. An approach to predicting the parameters of the methods for each testing database is proposed. Then a comparison is made between the tested methods and the state of the art.

Chapter 4, 'Retinal vessel quantification', describes the proposed approach to automation of AVR computation. An overview of the measures used for quantitative assessment of the retinal vessel structure is provided. The methods proposed for the automatic classification of the vessels into arteries and veins are reviewed. Then the proposed framework for estimation of the ratio is described. As a result, associations between the AVR and blood pressure of the subjects is assessed.

Chapter 5, 'Processing of the diaphragm image sequences', proposes a system for the automatic processing of diaphragm motion from dynamic MRI sequences. Diaphragm motion is separated into respiratory and non-respiratory motion. A set of features for the characterization of the motion is proposed, as well as a set of features characterizing the diaphragm's shape and position. The features are statistically compared between a group of normal subjects and a group of subjects with LBP. Lastly, an automatic approach to diaphragm motion detection, based on segmentation of the vessels in the retina, is proposed and validated on another set of measurements.

Chapter 6, 'Conclusions', discuss the achievements presented in the individual chapters and gives an overview of possible future improvements to the proposed frameworks.

Chapter II

Medical and technical background

Medical imaging covers, among other techniques, techniques for gathering quantitative information about the internal parts of the human body and organs, gathered both non-invasively and in vivo. These two important properties let the methods be employed in the diagnosis and research of pathologies that manifest themselves in living tissues. Many modalities are employed in order to visualize the various properties of the tissues, the most widely used visual examination methods include photography with the visible spectrum, multi-spectral imaging (which improves the spectral resolution and can use wavelengths that are invisible to the human eye, and can provide important information on the composition of the photographed object) and ultrasound-based imaging. In addition, there are methods based on advanced mathematical principles like tomographic approaches and MRI.

Naturally, the mentioned approaches are only a subset of the several proposed imaging modalities used to collect quantitative information on the organs. Even though the number of imaging methods is high, it is exceeded by the number of approaches proposed for processing the recorded images by image processing and analysis. A very broad and important field of the processing methods is the segmentation and characterization of structures in the medical images. The aim of the segmentation procedure is the delineation of the regions that are of interest – in a medical context, this typically means segmenting anatomical structures like organs, blood vessels and so on. The aim of the characterization procedure is then to provide a set of measures that can be used to depict the properties of the object of interest. The measures are then used to distinguish healthy and pathological structures, tissues and so on.

The segmentation and characterization techniques discussed here are those which are important for the context of the presented work. Two case studies are presented throughout the thesis which are focused on the characterization of the structures in different modalities. The first study is focused on characterizing the blood vessel structure in retinal photographs, and the second is oriented to characterizing the diaphragm and its motion in the body. Different motivations are behind the two studies. The retina is a vital organ with a double blood supply wherein numerous eye and systemic diseases manifest. At

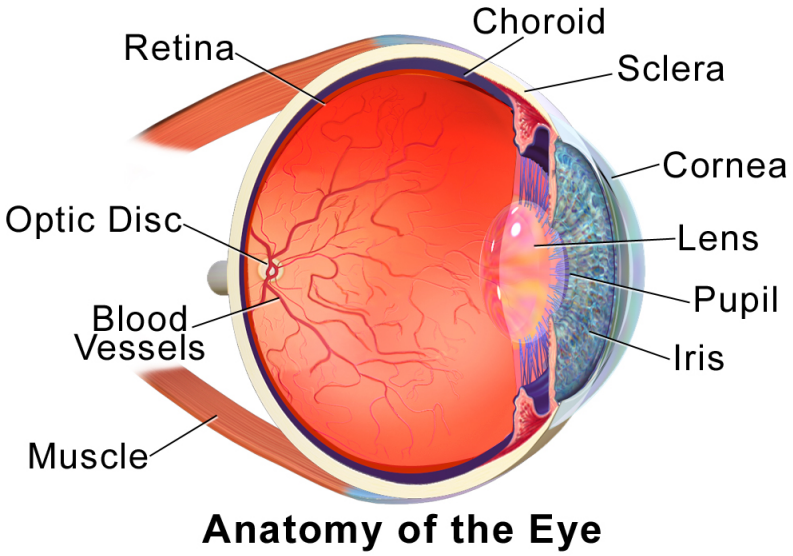


Figure 2.1: Anatomy of the eye – from [22].

the same time, it is an extension of the brain which allows for direct visual examination of the manifestation of diseases [4]. The diaphragm is the main respiratory muscle of the body and it also has an important role in body stabilization. Insufficient body stabilization can lead to problems like LBP. In the following chapters, the anatomical and physiological backgrounds of the retina and diaphragm are given.

2.1 The anatomy and physiology of the eye

Our eyes allow us to perceive 75–80 % of the environment around us. The principle of how the eye works is that it collects light and, through chemical reactions, changes the light energy into a neuronal signal which is processed in the visual cortex of our brain. The connection to the brain is important from the developmental point of view of the retina because the eye is basically an extension of our brain. Therefore, screening of the retina means direct in-vivo observation of brain tissue and, due to the blood supply, of our circulatory system [4].

The anatomy of the eye is depicted in Figure 2.1. The normally white eye ball is called the sclera, which has a transparent frontal part called the cornea. Under the cornea, there is the iris, which adjusts the amount of light entering the eye, and the lens, which focuses the light onto the back of the eye. The back part of the eye is where light-sensitive cells are located in a layered tissue called the retina. The retina itself is attached to the inner layer in the eye (the choroid) with the retinal pigment epithelium in the middle. The inner space of the eye is called the vitreous body and is filled with clear gel called the vitreous humour [21].

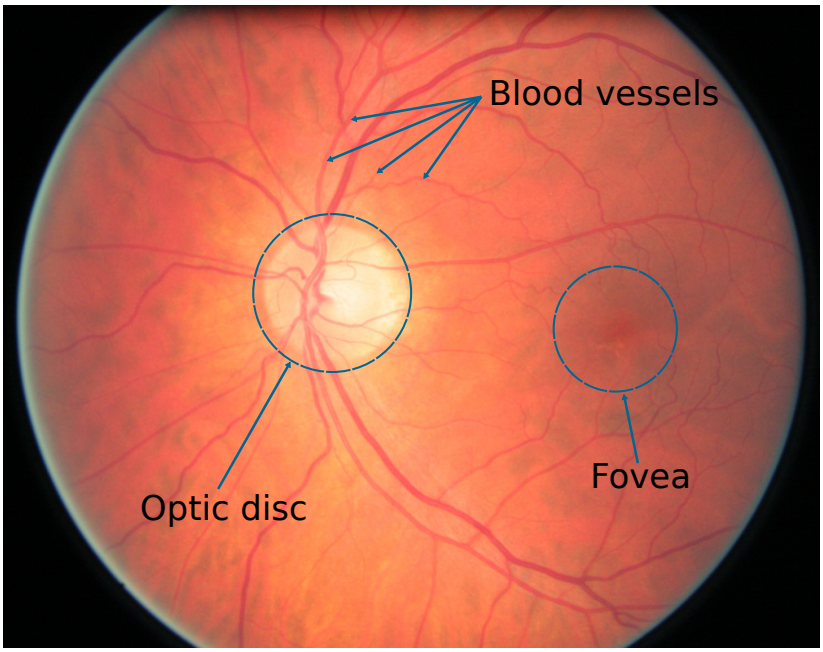


Figure 2.2: A photo of the retina with a depiction of the important anatomical parts.

The retina is the light-sensitive layer of the eye that is the most important anatomical part of the eye in the context of this work. The retina itself is a multi-layered tissue composed of different cells for light-energy conversion, the pre-processing of visual information and transmitting the neural signal. The photoreceptive layer is located furthest from the pupil, next to the choroid and pigment epithelium. The double blood supply is provided to the retina from the top and the bottom of the layer; the portion which comes through the choroid brings 65 % of the blood supply and the part coming from the top of the retina brings 35 %. The photoreceptive cells are divided into rods providing achromatic vision and cones providing colour vision [4].

The anatomy of the retina is depicted in Figure 2.2. The part responsible for pin-focus high-resolution colour viewing is the fovea, where the cones' density is the highest. On the rest of the retinal surface, rods outnumber the cones. The optic dics (OD) is the part of the retina where neuronal fibres and blood vessels enter the retina – no photoreceptive cells are located in the OD which is why it is also known as the blind spot. When the blood vessels enter retina inside the OD, one artery and one vein do so and then, by branching, they fill the retinal tissue. From a technical point of view, in the real three-dimensional space each vessel forms a tree-like structure with one root at the OD. In the retinal photographs, two-dimensional projections of the trees overlap, creating vessel crossings and cycles. However, an important property is that even in the two-dimensional projections, the arteries do not cross arteries and veins do not cross veins [23]. For an illustration of the differences between the arteries and veins, see Figure 2.4 in the Section 2.4.

From a diagnostic perspective, various diseases – including systemic diseases, eye diseases and diseases of the circulatory system – manifest themselves in the retina and provide observable and quantitatively measurable features for diagnosis [4]. The complications of such systemic diseases include diabetic retinopathy related to diabetes, hypertensive retinopathy from cardiovascular disease, and multiple sclerosis. As a consequence, the retina is vulnerable to organ-specific and systemic diseases. Imaging of the retina also allows diseases of the eye – as well as the complications of diabetes, hypertension and other cardiovascular diseases – to be detected, diagnosed and monitored.

Diseases manifesting themselves in the retina can be classified into diseases of the eye and systemic diseases. All of the following diseases belong to the group of the most common causes of blindness worldwide [4].

Diabetes mellitus is among the most prevalent diseases that manifest in the retina. There are approximately 150 to 200 million people with diabetes worldwide and 50 million in Europe alone [24]. The microvascular complication caused by diabetes in the retina is diabetic retinopathy.

Age-related macular degeneration (AMD) is another of diseases manifesting itself in the retina. The two main types are dry and wet AMD. Dry AMD, also called choroidal neovascularization, is the most threatening type for vision. It is accompanied by ingrowth of the choroidal vascular structure into the macula (the outer region around the fovea) and increased permeability of the vessels. The vascular ingrowth leads to rapidly deteriorating visual acuity, scarring of the pigment epithelium and permanent visual loss.

Glaucoma is a disease causing damage to the optic nerve and it also results in visual loss. The effect of the disease can be minimized by early detection and treatment. The changes brought about by glaucoma can be detected by using various types of retinal photographs and various types of measurements of the optic disc rim and its ratio to optic disc diameter is the important predictor of the disease.

Cardiovascular diseases, in the general sense, include all diseases of the vessels and heart. In a more particular sense it is used for diseases caused by atherosclerotic changes. Changes in the vessel structure can thus have an important role in the prediction and diagnosis of the diseases. Hypertension and atherosclerosis changes the ratio between the retinal arteries and veins (the AVR). Change in the AVR is also connected with the increased risk of stroke and myocardial infarction [4].

The segmentation and analysis of the blood vessel structure in the retina has an important role in the implementation of screening programs of several of the above-mentioned diseases [8, 25]: diabetic retinopathy, retinopathy of prematurity, arteriolar narrowing, hypertensive retinopathy, vessel diameter measurement in relation with diagnosis of hypertension. Furthermore, the vessel structure and its attributes serve in applications like foveal avascular region detection, computer assisted laser surgery, multimodal image registration, retinal image mosaic synthesis, optic disc identification and foveal localization [8].

2.2 Retinal quantification measures

The retinal quantification measures allow and have been used to describe the relationship between the systemic cardiovascular diseases and changes in the retina [26]. Associations

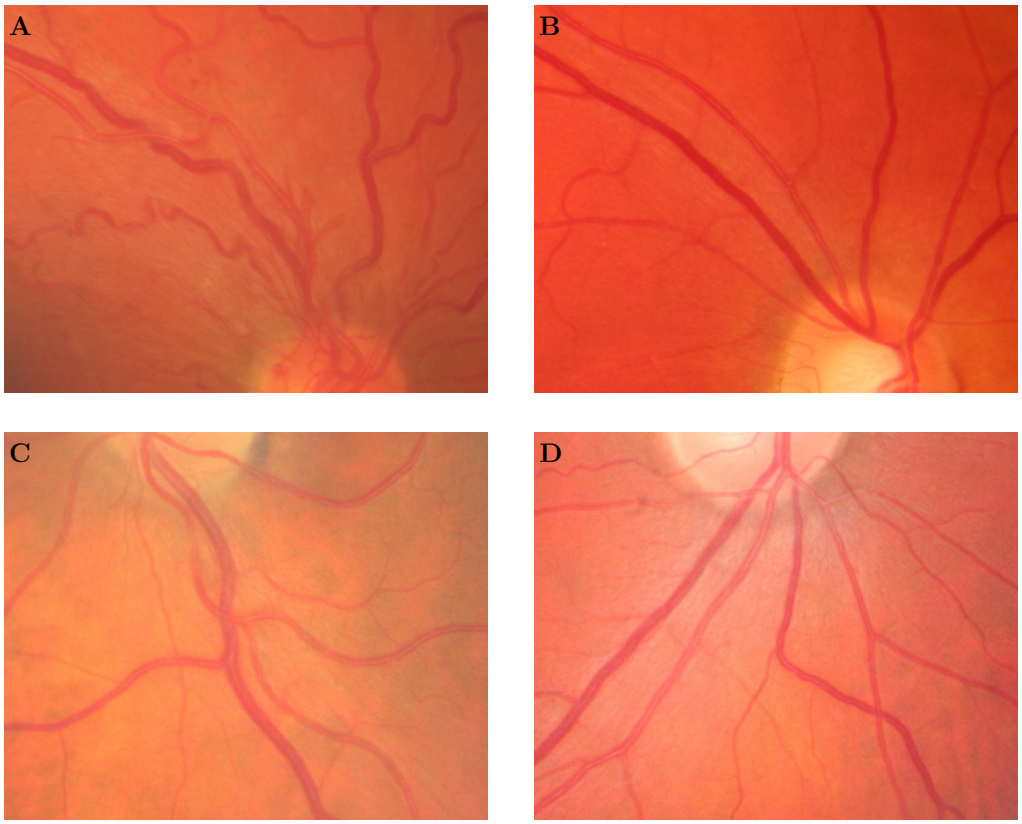


Figure 2.3: An illustration of how patient condition can be reflected in the retinal vasculature. On the image A, there is vasculature with venous beading which is typical for diabetic retinopathy and on the image B, there is healthy vasculature (HRF database, Section 2.6). On the image C, there is vasculature with narrow arteries, compared to the healthy vasculature on the image D. Arterial narrowing can be connected to various conditions like hypertension or atherosclerosis (Savitaipale database, Section 4.2).

between the measures and various clinical parameters such as age, blood pressure and body mass index (BMI) are being investigated in large population studies [26, 27] which allow for applying of the measures as predictors in diagnostic systems. The quantitative measures of the retina are all based on measurements of the blood vessel structure in the retina and include junctional exponents, angles at bifurcations, measures of vascular tortuosity, length:diameter ratios, fractal dimensions and AVR [26]. This section is devoted to the description of the various measures.

The junctional exponent refers to value of x in the equation $d_0^x = d_1^x + d_2^x$ which represents the diameters of the root vessel (d_0) with its branches (d_1, d_2). The theoretical value of the exponent approximate to value of 3 in healthy vascular networks in order to minimize power losses in the vascular structure [26]. Estimation of the junctional exponent can be troublesome in cases when the branches are bigger in diameter than the root vessel,

it is at the same time sensitive to diameter measurement errors. To overcome these limitations [28] developed a measure of deviation of the junctional exponent from its optimal value, $\rho = [d_0^3 - (d_1^3 + d_2^3)]^{1/3}/d_0$ which overcomes the aforementioned problems; also a significant difference of ρ between healthy subjects and subjects with peripheral vessel disease was concluded.

Vascular bifurcation angles are the angles between the branches at the vessel branching point. Theoretically, the optimal value for the angle has been estimated to be approximately 75°. The association of the angle with various clinical outcomes has been reported: reduced arteriolar angles were observed in hypertension, with increasing age and with low birth-weight males; a lower density of the vascular network was also observed to be connected with lower branching angles. No relation has been reported between bifurcation angles and peripheral vascular disease [26].

Vascular tortuosity (i.e. the measure of curvature of the vessels) has been used to measure the severity of conditions such as the retinopathy of prematurity (ROP), where increased arteriolar tortuosity belongs among the earliest predictors of plus disease. Venous beading has also been acknowledged as a feature of diabetic retinopathy.

The length:diameter ratio is calculated as the length from a particular vessel bifurcation to the midpoint in the preceding bifurcation and divided this by the diameter of the parent vessel at the bifurcation. It serves as a measure of the attenuation of the vessel and has been found to be increased in hypertension.

Fractal geometry of the vessel structure is used to assess its fractal dimension which, according to optimal junctional exponent $x = 3$, would be very close to 2 (optimal filling of the available space). Research estimated the fractal dimension to have the value 1.7 and arterioles to have a lower dimension than venules [26].

Arteriolar and venular diameters and their ratio, AVR, stand as the most widely used measures in the case of retinal vessel quantification. Typically the diameter of a vessel is estimated in the middle of the sides of the double-Gaussian cross-section profile which minimizes the defocusing effects of the image [29] on the diameter estimation. The main components of the AVR are the central retinal artery equivalent (CRAE) and the central retinal vein equivalent (CRVE) – estimates of the arteriolar or venular diameter as the vessels enter the retina through the OD. Those estimates are computed iteratively using veins and arteries within the area between 1 and 1.5 optic disc diameter (ODD)s from the OD's centre. Efforts have been made to research the formula in order to calculate the diameter equivalents. First Parr and Spears [30] proposed a polynomial equation for arteries:

$$W_c = \sqrt{0.87W_a^2 + 1.01W_b^2 - 0.22W_aW_b - 10.76},$$

Hubbard et al. later added a similar formula for veins [31]:

$$W_c = \sqrt{0.72W_a^2 + 0.91W_b^2 + 450.05},$$

where W_c is the diameter of the trunk vessel and W_a and W_b are diameters of the branches. These formulas expected the vessel to be paired iteratively in the way they branched producing CRAE and CRVE measures. The coefficients of the polynomials were estimated using measurements on samples of young normotensive subjects. This

approach was further improved by Hubbard in [32] where the iterative approach was disengaged in a way that vessels were combined in biggest–smallest pairs and, when there was an odd number of vessels, the remaining one was moved to the next iteration. The final rules of the approach consisted of using vessel branches when the diameter was $>80\mu m$ and vessels $<25\mu m$ were excluded [26]. So far, the last improvement in the way the AVR is computed has been proposed by Knudtson et al. [33] – the 'revised formula'. This important improvement took into consideration that the previously developed formulas need the diameters to be input in micrometres while in digital images the measurements are made in pixels. Another improvement was made regarding the fact that the number of measured vessels had a significant influence on the resulting AVR [33]. In [33] Knudtson et al. proposed using the six largest arterioles and six largest venules passing through the region of interest (ROI) (0.5–1 ODDs from the OD margin). The revised formula is based on the branching coefficient $b_c = \frac{w_1^2+w_2^2}{W^2}$, where w_i are the diameters of the branches and W is the diameter of the trunk vessel. Based on 44 healthy normotensive subjects, Knudtson et al. estimated the branching coefficient of arteries as 1.28 and of veins as 1.11. The diameter of the trunk vessel, based on those coefficients, is estimated for arterioles as

$$W = 0.88(w_1^2 + w_2^2) \tag{2.1}$$

and for venules as

$$W = 0.95(w_1^2 + w_2^2). \tag{2.2}$$

The CRAE and CRVE are then computed by iteratively combining vessels with the smallest and largest diameter from those six largest vessels. This approach simplified significantly the approach used previously by Parr and Spears [30] and by Hubbard [31]. Furthermore, Knudtson et al. reassessed several previously reported results using the revised formula and concluded the same association but with tighter confidence intervals. New studies have been employing the revised formula since [34, 35, 36].

The AVR was verified to be in significant relation to many factors – both systematic and ocular. The factors most notably include blood pressure, smoking, race, blood pressure, the risk of having a stroke, the risk of diabetes, BMI and age (for comprehensive list see [26]).

2.3 Retinal blood vessel segmentation methods

The segmentation of the blood vessels in the color fundus photographs is a relatively well understood and researched problem [4] and many methods have been proposed to solve the problem. In this section we review the blood vessel segmentation methods with a publicly available implementation. For a comprehensive review of the state of the art, see [8]. Examples of the segmentation output of the methods is presented in Figure 2.6.

The method proposed by Soares et al.[1] in [37] (the *Soares method*) is a supervised classification algorithm that uses the Morlet (Gabor) wavelet filter response as the classification feature. Three types of classifiers – the Gaussian mixture model (GMM), the k-nearest neigbours (k-NN) and least mean square error (LMSE) – are available for use. Furthermore, the green channel of an input image is by default added to the feature set.

[1] `https://sourceforge.net/projects/retinal/`

All features are normalized individually to zero mean and unit standard deviation (SD). The parameters of the method are: the set of Morlet scales (Λ_{mor}) used to define the classification features, the number of training samples (n_s), the number of Gaussians for modeling the vessels and non-vessels (n_{g1}, n_{g2}), and the number of iterations of the expectation maximization (EM) algorithm (n_i) used to define the GMM. The authors emphasize the efficiency of the Gabor transform in enhancing the vessel contrast. At the same time they conclude that there is a long training time, yet short classification time, for the GMM classifier. The simplicity of the implemented algorithms is emphasized. The reported disadvantages are false detections around high-contrast structures, like pathologies or the OD, and in rare occasions, bad tolerance of uneven illumination.

The method proposed by Sofka et al.[2] in [38] (the *Sofka method*) is a supervised classification algorithm based on multiscale matched filtering, and confidence and edge measures. The method extracts the vessel centrelines and, originally, its pixel-wise segmentation performance was not evaluated. However the method offers pixel-wise Likelihood Ratio Vesselness (LRV) as an output. The LRV measure, with subsequent binarization by thresholding was used in our experiments. The method is available as a pre-trained executable with no parameters to set. The authors claim a statistically significant improvement of the vessel segmentation performance over Frangi's vesselness measure and matched filter. The particular focus is on the detection of low-contrast and narrow vessels and the improvement of the classifier's performance on pathologies. The performance of the method is, however, assessed on thinned vessels due to the wider response of the filter around the vessel edges.

The method proposed by Azzopardi et al.[3] in [39] (the *Azzopardi method*) is an unsupervised algorithm that employs a bar-selective version of a 'combination of a shifted filter responses' (COSFIRE) filter – B-COSFIRE – which first filters the image with a difference of Gaussians (DoG) mask and then through the COSFIRE mechanism emphasizes the line patterns (creating response R). The final segmentation is obtained by thresholding. For the proper detection of vessel endings, the R of two types of endings (symmetric and asymmetric endings), are combined by averaging. Each R is defined by four parameters: the SD of the DoG filter (σ), the length of the line pattern (ρ) and two parameters allowing for spatial tolerance in the computation of R (σ_0, α). The authors emphasize the versatility of the COSFIRE filter as it can easily be rearranged to detect shapes other than the lines which were used in the case of vessel segmentation. The efficiency of the implementation and its robustness to noise are also emphasized.

The method proposed by Bankhead et al.[4] in [40] (the *Bankhead method*) is an unsupervised algorithm based on an isotropic undecimated wavelet transform (IUWT) [41] and binarization by percentile-computed threshold. After the binarization, post-processing by removing isolated objects and filling holes is done. The parameters are a set of wavelet levels (Λ_{ban}) for the IUWT, a percentile (p_t) used to compute the threshold value, and the sizes of the isolated objects and holes (ξ_s, ξ_h) for the post-process. The authors emphasize the high processing speed of the method and the simplicity of the setup, where the most important parameter – Λ_{ban} – has a small range of values even for images of very different resolutions. The disadvantage of the method is its slightly worse segmentation

[2] `https://www.cs.rpi.edu/~sofka/vessels_exec.html`
[3] `http://www.mathworks.com/matlabcentral/fileexchange/37395`
[4] `http://petebankhead.github.io/ARIA/`

performance when compared to the state-of-the-art methods.

The method proposed by Nguyen et al.[5] in [42] (*Nguyen method*) is an unsupervised algorithm based on line operators [43]. Vessel pixels are amplified by filtering the image with a mask of a defined size (W) that enhances pixels along lines with different orientations. Multiple filters with varying length of the line ($l_{1..n}$) together with the green channel of the input colour image are averaged to produce a single response with enhanced vessel contrast. The response is normalized to zero mean and unit SD. The number of filters is defined by step ω. The output of the algorithm is a gray-scale map. Thresholding (threshold τ) is used to produce the binary map. The authors emphasize the classification speed as an advantage of the method. Also, its local accuracy (segmentation near the vessel pixels) is claimed to be high. The method is supposed to handle well such areas that are often merged by other segmentation methods. The method is claimed to perform 'extremely well on non-pathological images' [42].

2.3.1 Vessel segmentation assessment methods

The performance of the blood vessel segmentation methods for the retina are usually assessed using measures where the binary segmentation output of a method is compared to the binary segmentation done by a human observer in a pixel-wise fashion. Accuracy (Acc), sensitivity (Sn), specificity (Sp) and area under the receiver-operating characteristic curve (AUC) are well established measures for the assessment [8]. Another measure – Matthew's correlation coefficient (MCC) – appeared recently in the vessel segmentation literature (for example, in [39]) and can give more insight into the evaluation when the sample sizes of the classes are skewed, which is the case in vessel segmentation. The performance measurement is typically done only on pixels inside the field of view (FOV), which is the circular region where the retinal surface appears. Throughout the presented work, the assessment of the segmentation methods was considered only on the FOV pixels.

2.4 Classification into arteries and veins

The vascular structure in the retina is physically cycle free (although its projection onto the 2D image plane becomes a vascular graph with cycles) [23]. One artery enters at the optic nerve head into the interior of the retina and branches without any reconnection; the same is true for veins.

Several features are of main interest when the vessels are manually classified into arteries and veins:

- Arteries are thinner, have a lighter red appearance and show a more clearly visible central vessel reflex than veins.
- At the crossings (in the 2D projection), only different vessel types are involved. In other words, an artery does not cross another artery and the same applies to veins.

The typical vessel structure close to the OD with delineated arteries and veins is depicted in Figure 2.4.

[5] `http://people.eng.unimelb.edu.au/thivun/projects/retinal_segmentation/`

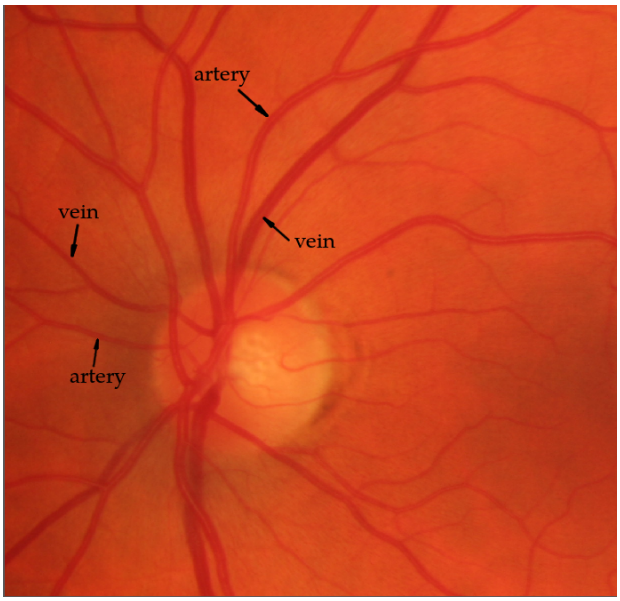


Figure 2.4: An example of arteries and veins close to the optic disc in RGB retina picture.

The automatic methods for vessel classification can be divided in two types: (i) approaches based on colour-based features and supervised or unsupervised classification and (ii) approaches employing colour-based features in combination with the underlying graph structure of the vessels. This thesis focuses on the feature-based classification of the vessels, thus, an overview of the methods proposed for the automatic classification of the retinal vessels into arteries and veins is presented.

Feature-based classification

Relan et al. experimented with both supervised [44] and unsupervised [45] classification. Prior to the feature extraction, the input image layers were normalized using the method of Chrástek [46]. Four classification features were used (in both supervised and unsupervised cases) – the mean of red, the mean of green, the mean of hue and the variance of red. The classification was pixel based, applied on the centreline pixels of the vessel segments, and the features were computed from a circular neighbourhood with diameter of $0.6 \cdot v_d$, where v_d is the vessel diameter around the pixels of interest. The least squares support vector machines (LS-SVM) classifier [47] was used in the supervised case and the GMM with EM fitting classifier [48] was used in the unsupervised case.

Grisan et al. [49] proposed an unsupervised method. The processed image was divided into four quadrants horizontally and vertically with the centre at the OD. The features – computed in a circular area around each vessel centreline point with diameter of $0.8 \cdot v_d$ – were the variance of red values and the mean hue value. Fuzzy C-means algorithm [50] was used for classification. Illumination and contrast was normalized by [51]. The approach was later enhanced by Tramontan et al. [52] by enhancing the tracing algorithm and

changing the AV classification scheme into a single feature – R contrast – which was computed from a vessel profile as a ratio between the peak value in a region around the central pixel and the higher of the edge values. The resulting values were fitted by the Hill function and classified by thresholding.

Kondermann et al. [53] first applied the 2D spline to the image layers and normalized the image illumination by subtracting the surface. Vessel profiles (vectors) and whole segments (matrices) were used directly as the features. The dimensionality of the features was reduced by multiclass principal component analysis (PCA) [54]. The support vector machines (SVM) and neural network (NN) classifiers were used for classification.

Saez et al. [55] experimented with various profile- and pixel-based features extracted from the red, green, hue of grey colour channels. Several different combinations of pixel- or profile-based features were tested: pixel values in the profiles from individual channels, pairs of pixels from green and red channels, a combination of the mean hue value and SD of the red value of a profile, the median value of a profile from each channel and the most frequent values among a channel, based on pixels or profiles. The median value of the green channel in a profile was chosen as the most discriminative feature. The k-means clustering was applied per-image to separate arteries from veins. The ROI was subsequently divided into quadrants and their rotation by 20° was undertaken to improve the classification by multiple overlapped clustering outcomes.

Niemeijer et al. published two papers dealing with vessel classification [56, 57]. In [56] the authors proposed 24 different features for arteriovenous (AV) classification including vessel width, vessel contrast, various averaged intensities and the second Gaussian derivatives of red, green, hue and saturation channels. Classification performance was assessed using linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), SVM and k-NN classifiers. It was concluded that k-NN had the best performance. The features were computed in every other centreline pixel of the vessels. Twelve features were selected for the classification. The approach was further expanded in [57] where 27 features were proposed consisting of the mean and SD of the vessel profile computed from hue, saturation, intensity, red and green channels and in red and green channels blurred by Gaussian with $\sigma = 2, 4, 8, 16$. The same classifiers used in the case of the previous study were tested, with LDA showing the best results. Soft labels assigned to each centreline pixel were transformed into segment labels by the median.

Muramatsu et al. proposed an approach based on LDA in [58]. The classification features were rather simple – the red, green and blue (RGB) values of the centerline pixel and the contrast of the RGB channels computed as the mean of the 5x5 region around centreline pixel which is subtracted from mean of a 10x10 region outside the vessel. The blue contrast feature was omitted, resulting in five features in the set used for the classification.

Dashtbozorg et al. proposed two methods for AV classification [59, 60]. The input images were preprocessed by the method proposed in [61]. In [59] Dashtbozorg et al. tested 30 different features based on RGB, and hue, saturation and value (HSV) image channels; the intensities of the centerline pixels; the mean and SD of the pixel intensities among a vessel segment; the maximum and minimum of the pixel intensities among a vessel; and the intensity of the centerline pixel in a Gaussian blurred channel (red and green only). Three classifiers were tested to do the AV classification: LDA, QDA, k-NN. The paper [60] proposes a simpler unsupervised approach to the classification when the vessel

Table 2.1: A summary of the features reviewed in Subsection 2.4. The letters R, G, B, H, S, V correspond to the image channels: red, green, blue, hue, saturation, value. Superscripts correspond to the papers where the features were employed: [1]Relan et al. [44] and [45], [2]Grisan et al. [49], [3]Tramontan et al. [52], [4]Kondermann et al. [53], [5]Saez et al. [55], [6]Niemeijer et al. [56] and [57], [7]–Muramatsu et al. [58], [8]Dashtbozorg et al. [59]

1. Variance in the centerline pixel neighborhood in R	[1]	[2]
2. Mean in the centerline pixel neighborhood in H		
3.-4. Mean in the centerline pixel neighborhood in R, G		
5. R contrast		[3]
6. Multiclass PCA of profile	[4]	
7. Multiclass PCA of rectangular vessel segment		
8. Median value of a profile in G		[5]
9.-13. Mean value of a profile in H, S, V, R, G	[6]	
14.-19. Intensity of the centerline pixel in H, S, V, R, G		[8]
19.-23. SD of of a profile in H, S, V, R, G		
24.-25. Highest intensity of a profile in R, G		
26.-27 Lowest intensity of a profile in R, G		
28.-35. Intensity of a centerline pixel in a Gaussian blurred (4 different sigmas) R and G channels		[8]
36. Intensity of the centerline pixel in B	[8]	[7]
37.-39. R, G, B contrast (5x5 inside / 10x10 outside)		
40.-43. Mean intensity of the vessel in R, G, B, H, S, V	[8]	
43.-46. SD of the intensity among the vessel in R, G, B, H, S, V		
47. Max intensity among the vessel in R		
48. Min intensity among the vessel in R		

pixels of the red channel (after normalization of the image) are considered and clustered (using k-means) into artery, vein and *unknown* clusters.

An overview of all the reviewed features can be found in Table 2.1.

2.5 Methods for automatic estimation of arterio-venous ratio

Several systems for the automation of the AVR estimation procedure were proposed. The general steps, followed by the both manual and automatic approaches, can be summarized as:

1. Image preprocessing
2. OD localization and estimating the ROI
3. Vessel segmentation
4. Cutting the vessels into segments and estimating the vessel width
5. Estimating the sub-trees of the vessels

6. Classifying the vessels as arteries and veins
7. Selecting the final vessels or vessel pairs for estimation of CRAE and CRVE
8. Computing the AVR

Not all these steps have to be present. The final vessel selection differs typically between the manual method to AVR estimation and the automatic methods. Selecting the final vessels for AVR computation in the manual approaches is typically done by selecting either the root part or branches of a vessel. The automatic approaches typically do the computation in sub-regions of the ROI (e.g. by division into six sub-rings) and assess the AVR separately in each sub-region. The final AVR value is averaged over the sub-regions [57, 60]. So far the validation of the automated AVR systems has only been made by direct comparison of the automatically and manually obtained AVR. No association between blood pressure or other clinical measurements has been investigated, although it has been concluded as a desired progress [57]. A review of the proposed systems for fully automatic AVR estimation follows.

Tramontan et al. [52] proposed a system where the image illumination is normalized, both intra- and inter- image. The algorithm uses vessel tracing to segment the vessel network that is subsequently used to detect and segment the OD. It was not specified how the final vessels are selected for estimating the central vessel equivalents. The proposed system was evaluated on a non-public dataset. Evaluation of the system was done by investigating the correlation between the manual and automatic values.

Niemeijer et al. [57] proposed an approach which employed the vessel classification procedure reviewed in Subsection 2.4. For the vessel width measurement, a technique called 'tobogganing' [62, 63] was used in the both the manual and automatic approaches. The INSIPRE-AVR (see Subsection 2.5.1) dataset was used for evaluation of the system. The manual and automatic AVR values were compared using Bland-Altman plots and Student's t-test.

Dashtbozog et al. [60] proposed a system based on the vessel classification approach described in Subsection 2.4. A scheme for vessel graph estimation was proposed and used to improve the vessel classification performance. The system was validated using the DRIVE, INSPIRE-AVR and VICAVR databases.

2.5.1 Assessment methods for arterio-venous ratio

To assess the performance of AV classification, most of the methods used the classification accuracy of centerline pixels [44, 45, 53, 58, 57], others assessed classification accuracy of the whole segments [49, 55]. The assessment is typically limited to the segments that are within or around the ROI for the AVR computation.

To assess the whole AVR framework, the correlation coefficient between the manually and automatically estimated AVR is often used. However, the correlation coefficient has drawbacks because the manually estimated AVR can be biased, which is why associations of the AVR with the blood pressure are used. The association of the AVR with blood pressure is well documented [27], thus it serves as an objective way to validate the proposed framework. The assessment is done typically in a way that a linear model is

estimated between the AVR and blood pressure of the measured patients and confidence intervals and slope of the linear regression are assessed and compared [27]. This way of assessment of an automatic AVR framework, has not been done yet, although it has been considered as a desired outcome [57].

2.6 Databases used in retinal image processing

Medical image databases with an appropriate GT about the image contents enable the development and proper evaluation of automatic image analysis methods. In this section, databases containing GT for the blood vessels and databases containing GT for the estimation of AVR are reviewed.

Five publicly available databases were identified, which have the GT for the blood vessels: ARIADB, CHASEDB1, DRIVE, HRF and STARE. Information about the databases selected for testing the retinal blood vessel segmentation methods is summarized in Table 2.2. The number of images, image dimensions, FOV angle and diameter, the subsets and percentage of vessel pixels in the GT are presented, for further information please refer to the original publications. Examples of the database images are provided in Figure 2.5.

Table 2.2: A summary of database information. N_i is the number of images and N_{GT} is the number of experts and the percentage of annotated vessel pixels (per expert) in the ground truth segmentation. Abbreviations of the image subsets are age-related macular degeneration (AMD) and diabetic retinopathy (DR).

Name, ref.	N_i	FOV [°]	Dimensions FOV ⌀	Subsets	N_{GT}
ARIADB [64]	143	50°	768x576 739 px	AMD (23) Healthy (61) DR (59)	2 (9.6%, 8.5%)
CHASEDB1 [65]	28	30°	999x960 916 px	Left eye (14) Right eye (14)	2 (10.1%, 9.7%)
DRIVE [66]	40	45°	565x584 540 px	Training (20) Test (20)	2 (12.7%, 12.3%)
HRF [67]	45	60°	3504x2336 3262 px	Healthy (15) DR (15) Glaucoma (15)	1 (9.13%)
STARE [68]	20	35°	700x605 649 px	–	2 (10.3%, 14.8%)

The retinal databases which are publicly available with the GT for the AVR are summarized in the following paragraph. Only a short overview with the basic information is given here to provide background information, as none of the databases were used in our experiments with the AVR. Those databases include DRIVE [66], which contains also

the GT for the blood vessels and is covered at the beginning of this section; INSPIRE-AVR [57], which contains 40 colour images of the vessels, ODs and an AVR reference standard – the image resolution is 2392x2048; VICAVR [69], which contains 58 images which include the calibre of the vessels measured at different radii from the OD as well as the vessel types labelled by three experts – the image resolution is 768x584; WIDE [70], which contains 30 high-resolution (although downsampled), wide-field images of healthy eyes and eyes containing AMD – the image resolution differs with an average of around 800 x 1200. Other, non-public databases can be also found [44, 49].

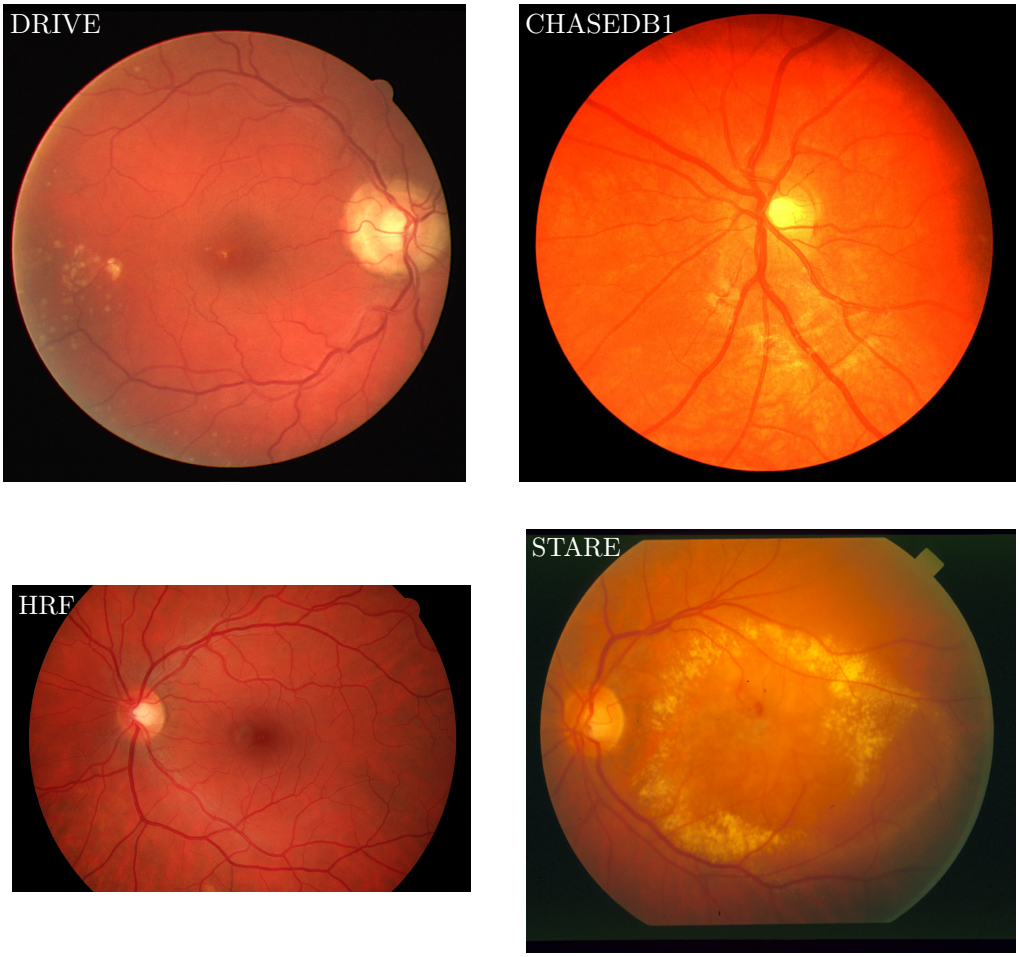


Figure 2.5: Example images from the databases used for the assessment of the blood vessel segmentation methods. The images are labelled by the names of the databases. It can be seen that the retinal images have similar characteristics but can differ in color, contrast, texture and presence of pathologies (clearly visible brighter dots in DRIVE and STARE images for example). The pathologies, OD and central vessel reflex (a brighter strip typically visible in arteries) are the most common sources of misclassifications in the vessel segmentation process [8].

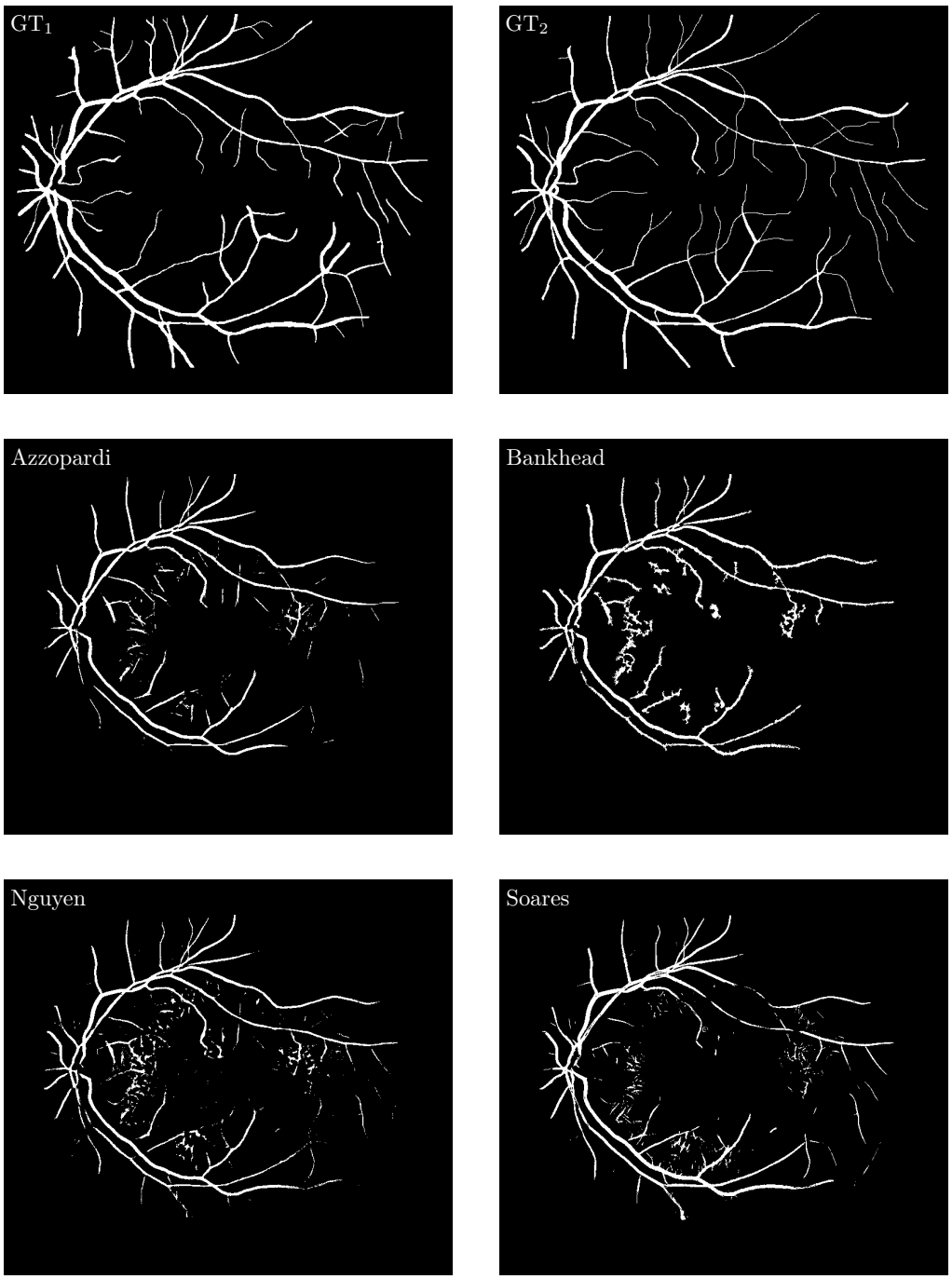


Figure 2.6: Examples of the binary blood vessel segmentation (optimized for the segmentation Acc with GT_1) of the STARE image from Figure 2.5. Four methods out of the five reviewed methods – those that give better results – are illustrated. It can be noted how the pathologies cause false positive detections. Also differences between the manual segmentations GT_1 and GT_2 are clearly illustrated.

2.7 The anatomy and physiology of the diaphragm

The diaphragm is the main breathing muscle. It is 2–4 mm thick and located at the bottom of the thoracic cavity, as illustrated in Figure 2.7. The diaphragm is not actually a single muscle but rather a composition of several parts which can be activated separately. It has a concave, asymmetric shape with the most cephalic point – the centrum tendineum – connecting the parts together. The diaphragm separates the thoracic and abdominal walls and the abdominal cavity [71].

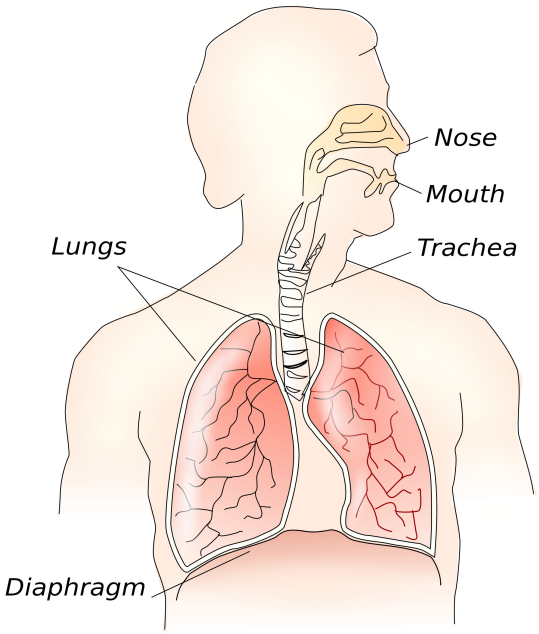


Figure 2.7: An illustration of the diaphragm's location inside the trunk. Picture by Theresa Knott, distributed under a CC-BY 3.0 license – taken from [72].

The diaphragm and deep stabilization muscles of the body have been described as an important functional unit for dynamic spinal stabilization [73, 74]. The diaphragm precedes any movement of the body by lowering itself and subsequently establishing abdominal pressure, which helps to stabilize the lumbar part of the spine. Proper activation of the diaphragm within the stabilization mechanism requires the lower ribs to be in an expiratory (low) position. During the breathing cycle, the lower ribs have to stay in the expiratory position and only expand to the sides. This is an important prerequisite for

a straight and stabilized spine. Under these conditions, the motion of the diaphragm is smooth during respiration and properly helps to maintain the abdominal pressure.

Dysfunction of the cooperation among the diaphragm, abdominal muscles, pelvic floor muscles and the deep back muscles is the main cause of vertebrogenic diseases and structural spine findings, such as hernia, spondylosis and spondylarthrosis [75, 76]. Diaphragm function control is a broad and important issue for a number of fields of investigation, including pulmonology [77], chest surgery, rehabilitation [78] and gastroenterology [79]. However, studies dealing with the lumbar stabilization system mostly do not include monitoring diaphragm activity [80]. A traditional objective of studies dealing with the diaphragm's function is to study the diaphragm's respiratory function [81]; studies focused on the postural function are rare.

Studies focused on diaphragm activation with the aim of posture stabilization include those of Hodges [82, 83, 84, 85], who concluded that phase modulation corresponds to the movement of the upper limbs in diaphragm electromyography records. Some works deal with various modes of the diaphragm's functions in various respiration types [86, 87] or in situations not directly related to respiration, like activation during breath holding [88]. These studies have always concentrated on healthy subjects who did not exhibit symptoms of respiratory disease or vertebrogenic problems.

The causes of LBP and their relations to spinal findings have been the subject of several studies and continue to be a significant research topic. Jensen [89] assessed low back MRI with the goal of finding structural changes related to LBP. Jensen found no direct connection between certain types of structural changes and LBP. The only structural change related to pain was disk protrusion. Carragee [90] studied the MRI findings of 200 subjects after a period of LBP and found no direct and significant MRI finding related to LBP.

The way in which the diaphragm is used for non-breathing purposes is affected by its recruitment for respiration [91]. There is evidence that the presence of respiratory disease is a stronger predictor for LBP than other established factors [11]. However, the relationship between the respiratory function and the postural function is widely disregarded [14]. The coordination of muscles in the body for posture stabilization is a complex issue, and the role of the diaphragm in this cooperation has not been intensively studied [13].

2.8 Diaphragm assessment using magnetic resonance imaging

MRI has not been a modality used very often in the studies of the diaphragm. Although several studies exist addressing suitability of the MRI images for the measurements or assessing the influence of image artifacts. A review of the existing studies that employed MRI in the assessment of the diaphragm is provided in this section. Example of the diaphragm screened using MRI is provided in Figure 2.8.

Gierada [92] used MRI for observing the anteroposterior (AP) size of the thorax, the height of the diaphragm during inspiration and expiration, and also the ventral and dorsal costophrenic angle during maximal breaths in and out. Gierada used a 1.5 T MRI device for measuring the height of the excursions of the diaphragm at three different points in several sagittal planes. Gierada [93] assessed MRI artifacts and concluded that MRI is a valid method for diaphragm image processing along the diaphragm contour.

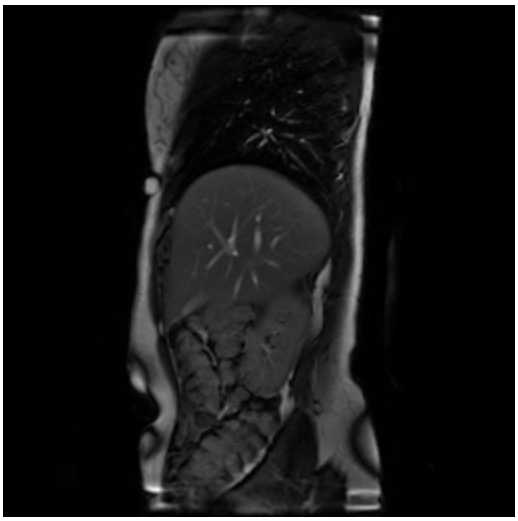


Figure 2.8: An example of the diaphragm screened by MRI in a sagittal plane.

Differences in diaphragm movement while performing thoracic or pulmonary breathing with the same spirometric parameters were tested by Plathow [94]. Plathow also examined the vital capacity (VC) of the lungs compared with 2D and 3D views in [95]. He concluded that there was a better correlation between lung capacity and the 3D scans. In another study, Plathow focused on the dynamic MRI. He proved significant correlations among diaphragm length and spirometric values vital capacity, forced expiratory volume (FEV1) and other lung parameters [96]. Plathow [97] assessed diaphragm length using dynamic MRI in the mid-coronal plane by 1.5 T magnetic resonance, and concluded that the spatial and time resolution was sufficient for acquiring the breathing sequences.

Suga [98] used breathing MRI (BMRI) for comparing healthy subjects and subjects with chronic obstructive pulmonary disease (COPD), measuring measuring the height, excursions and antero-posterior (AP) size of the diaphragm with the zone of apposition. Suga concluded that BMRI is a useful non-invasive method with good spatial and temporal resolution.

Kotani [99] and Chu [100] assessed chest and diaphragm movements for scoliosis patients using breathing MRI. Kotani [99] concluded that there was ordinary diaphragm motion with limited rib cage motion in the scoliosis group. The position of the diaphragm was measured relative to the apex of the lungs to the highest point of the diaphragm. Chu [100] compared healthy subjects against subjects with scoliosis, finding the same amount of diaphragm movement in both groups. The scoliosis group had the diaphragm significantly lower in the trunk and relatively smaller lung volumes.

Paradox diaphragm movements for subjects with COPD were investigated by Iwasawa [81] using dynamic MRI. Iwasawa used deep breath sequences while comparing the diaphragm height and costophrenic angles. Iwasawa also used a displacement area to measure movement of the diaphragm.

The effect of intra-abdominal pressure on the lumbar part of the spine was observed by

MRI and pressure measurement by Daggfeldt and Thorstensson [101].

The measured parameters can be summarized as follows:

- Measurements of cranio-caudal excursions of the diaphragm [102, 103, 98, 104, 92, 105]
- The anteroposterior and lateral proportion of the diaphragm [99, 100]
- The height and anteroposterior proportion of the diaphragm [92, 81]
- Shortening of the diaphragm contour in the sagittal and frontal plane [106]
- Curvature of the diaphragm [107]
- Comparison between the movement of the ventral and dorsal diaphragm part [93]
- Assessment of the diaphragm motion using displacement area of the diaphragm [81, 88]

2.9 Relation between low back pain and diaphragm

LBP is a wide-spread and widely studied phenomenon. Alternating respiratory patterns and anatomical changes in the diaphragm have been assessed in LBP subjects. Studies concluding that there is increased susceptibility to pain and injury [73, 84, 108] identified differences in muscle recruitment in people suffering from LBP.

Janssens [15] used the fatigue of inspiratory muscles and observed an altered postural stabilizing strategy in healthy subjects. Jenssens also observed non-worsening stabilization with an already altered stabilizing strategy in subjects suffering from LBP.

Kolar [109] investigated differences in diaphragm contractions between healthy subjects and LBP subjects. He observed lesser contractions in the posterior part of the diaphragm while the postural demands on the lower limbs increased, and he suspected that the intra-abdominal pressure lowering might be the underlying mechanism of LBP.

Grimstone [12] measured respiration-related body imbalance in subjects suffering from LBP, observing worse stability in subjects with LBP. Roussel [14] assessed the altered breathing patterns of LBP subjects during lumbopelvic motor control tests, concluding that some subjects used an altered breathing pattern to provide stronger support for spinal stability.

Chapter III

Retinal blood vessel segmentation

3.1 Introduction

In this chapter, the performance of pixel-wise blood vessel segmentation methods designed for retinal images with publicly available implementation is reviewed and compared. Five methods, two supervised and three unsupervised, were studied, and their performance was assessed using five publicly available databases with GT for vessel segmentation.

To assess the performance potential of each method, an experiment was set up so that the parameter space of each method was sampled for each database and the optimal segmentation performance was searched using a grid search (see Subsection 3.3.1). In order to keep the conditions of the experiment consistent, the image preprocessing parts of the methods were separated and applied to all the methods.

Analysis of the algorithms' performance over the parameter value ranges provides several valuable results: (1) a comparison and assessment of the generalization capability of the methods over several different databases that exploits the segmentation potential of each method, (2) a comparison of the performance and settings of the methods with the original publications and with the state of the art, and (3) baseline settings suitable for application of the methods to other data.

Results from the parameter optimization were then used to propose a set of models which allow to predict the parameters of the segmentation methods for new databases. The prediction is based on various features of the databases like image resolution or OD size.

3.2 Data

The databases used to assess the segmentation performance and optimize the methods' parameters were reviewed in Section 2.6. The databases are ARIADB, CHASEDB1, DRIVE, HRF and STARE.

The DRIVE (Digital Retinal Images for Vessel Extraction) and STARE (Structured Analysis of the Retina) databases are currently the datasets most commonly used to evaluate retinal vessel segmentation methods. All the methods reviewed in this chapter were originally evaluated using DRIVE and STARE, with an exception of the *Bankhead method*, which was only evaluated using DRIVE. Our experiments were done on three other databases with different properties, which enabled a more robust comparison of the methods and allowed modeling and predicting the settings of the methods' parameters.

3.3 Methods

The methods reviewed for the blood vessel segmentation task were introduced in Chapter 2, Section 2.3. The methods are the *Azzopardi method*, *Bankhead method*, *Nguyen method*, *Soares method* and *Sofka method*.

3.3.1 Performance optimization and comparison

The segmentation performance of each algorithm was assessed using the manual segmentation of a database's first observer (sorted in alphabetic order) as a reference and using MCC, Acc, Sn, Sp and AUC as the performance measures. In the case of databases without a FOV mask (ARIADB, CHASEDB1, STARE), the mask was generated using edge detection and ellipse fitting. The measures are expressed as percentages. If two sets of manual segmentations were provided in a database, the performance of the second observer (the second manual segmentation) was assessed and compared to the performance of the automatic methods.

The setup of the grid search

A grid search was used to find the best segmentation performance in a subspace created by a Cartesian product over the sampled sets of the parameter values in Table 3.1. The parameters of the methods were optimized for highest Acc and AUC[1]. Below we discuss the details for each method that differ from standard grid search procedure.

With the *Soares method*, the parameter settings of the classifiers are determined by preliminary experiments and the reported values were set as indicated in Table 3.1. To search for the best Λ_{mor}, a greedy optimization approach was used in order to avoid evaluating poorly performing Λ_{mor}.

With CHASEDB1 and HRF, it was infeasible to execute the *Sofka method* on the images in the original resolution. Therefore, the images were downscaled so that the longer side of an image was 600 pixels.

The *Azzopardi method* is set up using eight parameters, which is a relatively large number. As a consequence, an exhaustive search in the parameter space requires evaluating the performance at too many points. Therefore, it was decided to explore the performance for the parameters individually beginning at the point defined in [39] or at a randomly

[1]Complete results from the optimization procedure are available at `http://www.it.lut.fi/mvpr/medimg`

Table 3.1: Sampled values of the parameters. The superscript over the set of wavelet levels (e.g., $\{1,2\}^{\leq 2}$) represents subsets of the indicated size (e.g. $\{1,2\}^{\leq 2} = \{\{1\},\{2\},\{1,2\}\}$). For the *Azzopardi method* assessment of the algorithm performance was done separately for individual parameters beginning at the 'starting point' and adding values indicated in the column Test ranges. For the *Bankhead method* ξ_s was sampled logarithmically in the interval $\langle 0, \text{max value (all vessels removed)}\rangle$ on each database. AR, CH, DR, HR, ST are abbreviated names of the databases which are ARIA, CHASEDB1, DRIVE, HRF, STARE.

Soares method		*Bankhead method*		*Nguyen method*	
Λ_{mor}	$\{1, 2\cdots 17\}^{\leq 3}$	Λ_{ban}	$\{1,2,\cdots,4\}^{\leq 4}$ (AR,CH,DR,ST)	W	$\{12,13,\cdots,20\}$ (AR, DR, ST)
n_s	$\{2\cdot 10^5\}$ (AR, DR, ST)		$\{1,2,\cdots,5\}^{\leq 3}$ (HR)		$\{45,46,\cdots,55\}$ (HR)
	$\{3\cdot 10^5\}$ (CH, HR)	p_t	$\{0.05, 0.1, \cdots, 0.3\}$		$\{25,26,\cdots,35\}$ (CH)
n_{g1}	$\{30\}$	ξ_s	40 values in log. scale	ω	$\{2,4,\cdots,W-1\}$
n_{g2}	$\{40\}$	ξ_h	$\{0\}$	τ	$\{0.5, 0.55, \cdots, 1.5\}$

Sofka method		*Azzopardi method*						
τ, pad	$\{0.50, 0.52, \cdots, 2.0\}$ (AR)		Starting point					
preprocessing	$\{0.50, 0.52, \cdots, 1.5\}$ (CH,DR)	Parameter	AR	CH	DR	HR	ST	Test ranges
	$\{0.50, 0.52, \cdots, 1.5\}$ (HR,ST)	σ_1	2.5	4.8	2.4	7.2	2.7	$\{0, \pm 0.2, \pm 0.5\}$
τ, CLAHE	$\{20.0, 20.02, \cdots, 21.0\}$ (AR)	σ_2	2	4.3	1.8	6.8	2.1	$\{0, \pm 0.2, \pm 0.5\}$
preprocessing	$\{18.0, 18.02, \cdots, 20.5\}$ (CH)	r_1	10	18	8	26	12	$\{0, \pm 4, \pm 6\}$
	$\{8.0, 8.02, \cdots, 14.0\}$ (DR)	r_2	24	34	22	50	24	$\{0, \pm 4, \pm 6\}$
	$\{31.50, 31.52, \cdots, 33.5\}$ (HR)	σ_{01}	2	3	3	2	1	$\{0, \pm 1, \pm 0.5\}$
	$\{20.0, 20.02, \cdots, 22.0\}$ (ST)	σ_{02}	1	1	2	1	1	$\{0, \pm 1, \pm 0.5\}$
		a_1	0.4	0.2	0.7	0.4	0.6	$\{0, \pm 0.5, \pm 0.2\}$
		a_2	0.1	0.1	0.1	0.1	0.1	$\{0, \pm 0.5, \pm 0.2\}$
		τ	$\{0.1, 0.105, \cdots, 0.2\}$					

selected point. Then, the ranges defined in Table 3.1 were followed from the starting point only in the direction of the parameter axes. The whole procedure was repeated from the best observed point as long as there was an improvement in the performance.

The performance of the *Bankhead method* was observed to be influenced very little by parameter ξ_{h}, and thus, ξ_{h} was not used in the experiment. It should be noted that the AUC measure was assessed on the vessel-enhanced image before binarization and cleaning as the post-processing step. As a consequence, it yielded low values. It is possible to observe the influence of the cleaning on the receiver-operating characteristic (ROC) characteristics in Figures 3.3 – 3.5.

Training data

Subsets of the databases were used as training data to optimize the parameters and to train the classifier of the *Soares method*. The number of training images differed for each database: with DRIVE, the dedicated training set was used. With HRF, 15 random images were selected for training. With ARIADB, 30 random images were selected. With STARE and CHASEDB1, each database was randomly divided into two subsets of the same length and each subset was used to train a classifier. Each of the classifiers was then used to classify images from the training set of the other classifier.

Image preprocessing

Two different preprocessing approaches of the input images were identified among the methods. The first one is the 'pad only' preprocessing method, which pads the edges of the FOV [37] and forms a part of the *Soares method*. The second one is the contrast limited adaptive histogram equalization (CLAHE) preprocessing method, where the image is padded as above and then CLAHE is applied. This approach comes as a part of the implementation of the Azzopardi method. With the pad-only preprocessing, padding by 50 pixels was used. With the CLAHE preprocessing, the image was padded to the image edges before applying the CLAHE algorithm with 6x6 tiles per image.

Each method except the *Sofka method* was applied to the green channel of the input image. The *Sofka method* can be applied to the green channel alone or to the full colour image. It produced better results when applied to the colour image.

3.3.2 Prediction of the segmentation parameters

Taking advantage of having multiple test databases with different image resolutions, we aimed to fit linear models in order to be able to predict the settings of the methods for new databases. Eight features of the databases were chosen and tested as a predictors. Linear models were then established based on various subsets of the predictors and assessed using statistical properties of the models.

The established predictors are in Table 3.2, d_r is the angular resolution of the segmented database expressed in pixels per FOV angle, d_n is the percentage of the vessel pixels in the vessel , d_{od} is the mean diameter of the OD, d_{nf} is the number of pixels in the FOV mask and d_{df} is the diameter of the FOV mask in pixels. Lastly $d_{v1..3}$ corresponds to

Table 3.2: A description of the predictors used to model the parameters of the segmentation methods. Values marked * refer to the values from Table 2.2. A description of the computation of vessel width is provided in the text.

Predictor	Definition
d_r	$\frac{\text{FOV}\varnothing}{\text{FOV}[°]}$ *
d_n	Percentage of the vessel pixels in the manual segmentations N_1*
d_{od}	Mean optic disc diameter
d_{df}	$FOV\varnothing$ in pixels.
d_{nf}	Number of pixels in the FOV.
$d_{v1..3}$	Mean vessel width

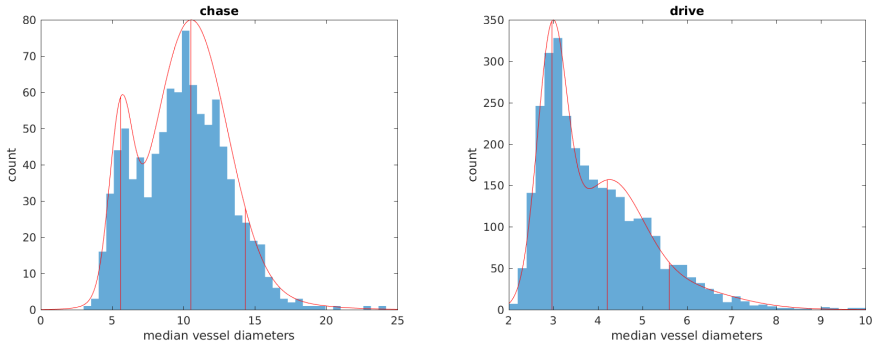


Figure 3.1: An example of the histogram of vessel width estimated on different databases. The means of the Gaussians estimated by GMMs were used as the predictors of the vessel segmentation parameters.

vessel width. The *Bankhead method* was applied on the retinal images which resulted in preliminary vessel segmentation. The *Bankhead method* was selected because its parameters are stable with varying image resolution. The segmented vessels were also separated into vessel segments using Bankhead's toolbox [40]. A length limit was applied to the resulting vessel segments so only reasonably long segments were processed. The statistical distribution of the vessel widths was then fitted with three Gaussians by using the GMM. The mean of each model component was then used as the predictor in ascending order (d_{v1} corresponded to the lowest μ, d_{v2} to the second lowest and so on). The distribution of the vessel widths is illustrated in Figure 3.1.

The predictors were used to estimate linear models and adjusted R^2 was assessed for each model to determine its ability to predict the segmentation parameters. Combinations of one to three predictors were used. The segmentation parameters were divided into two groups – a resolution-related group and all the others were grouped. The resolution-related group was expected to be predictable via the selected predictors and, thus, a model was searched for in order to give as good as possible prediction for all parameters in the group. Application of the models on the whole group was also done in order to be able to draw conclusions about the generalizability of the models and decrease the

Table 3.3: The division of the parameters on the resolution-related group and the non-resolution-related group.

Category	Parameters	
Resolution-related	Azzopardi	$\sigma_{1..2}$, $r_{1..2}$
	Bankhead	$\Lambda_{ban1..2}$, ξ_s
	Nguyen	W
	Soares	$\Lambda_{mor1..3}$
Other	Azzopardi	$\sigma_{01..02}$, $a_{1..2}$, τ
	Bankhead	p_t
	Nguyen	ω, τ
	Soares	—

influence of random perturbations in both predictors and parameters. The parameters in the 'non-resolution-related' group were inspected separately. The inspection was done for the parameters optimized by both AUC and Acc measures.

3.4 Results

The results from the optimization experiment and the comparison of the results with the state-of-the-art methods are reported in Subsection 3.4.1. Then Subsection 3.4.2 reports parameters corresponding to the results reported in Subsection 3.4.1 and deals with the prediction of the parameters.

3.4.1 Performance of the algorithms

The performance of the algorithms is presented in Table 3.4. It is possible to refer to Figure 3.2 for visual comparison of the resulting accuracy and its change with change to the preprocessing. The ROC characteristics of the methods are provided in Figures 3.3 – 3.5 for a more general comparison of the methods. In the ROC comparison, each method is presented only with the preprocessing type that resulted in better performance, applied to all databases, in order to keep the amount of data reasonable.

First we compare the influence of preprocessing on each method and focus only on the preprocessing approach that leads to better performance. Influence of the preprocessing selection is illustrated in Figure 3.2. Applying CLAHE preprocessing had a positive effect on the *Azzopardi and Bankhead methods*. With the *Nguyen method*, the effect was positive on average, but the results were inconsistent. In general, the choice of the preprocessing approach had the smallest effect on the *Nguyen method*. The effect on the *Soares method* was negative, with the exception of the STARE database. The effect on the *Sofka method* was negative for all databases. The absolute difference between the accuracy measured with CLAHE and pad only was up to 0.5 in percentage units. As a result, the comparison that follows will consider the results from CLAHE preprocessed

Table 3.4: The performance of the methods, each assessed with selected type of preprocessing: the *Soares and Sofka methods* with pad only and the *Azzopardi, Bankhead and Nguyen methods* with CLAHE. Corresponding parameter settings are given in Subsection 3.4.2. The results published in the original papers are shown in italics under the current results. The best scoring results in sense of Acc and AUC for each database are shown in bold.

Acc optimized; Pad only – Soares, Sofka. CLAHE – Azzopardi, Bankhead, Nguyen.

	Soares method					*Sofka method*					*Azzopardi method*				
—	Acc	Sn	Sp	MCC	AUC	Acc	Sn	Sp	MCC	AUC	Acc	Sn	Sp	MCC	AUC
AR	93.6	53.6	97.7	57.6	**90.7**	93.3	44.5	98.3	53.3	86.3	**94.0**	56.2	97.9	60.8	89.2
CH	**94.6**	69.0	97.4	68.9	**96.4**	93.2	50.9	98.0	58.0	90.8	94.3	63.7	97.8	66.7	93.2
	—	—	—	—	—	—	—	—	—	—	*93.9*	*75.9*	*95.9*	*68.0*	*94.9*
DR	**94.7**	71.7	98.1	75.0	**96.1**	93.5	60.9	98.2	67.9	91.5	94.5	70.0	98.1	74.0	95.6
	94.7	—	—	—	—	—	—	—	—	—	*94.4*	*76.6*	*97.0*	*74.8*	*96.1*
HR	95.8	73.4	98.0	73.3	**97.0**	94.3	58.3	97.8	61.4	93.7	95.7	69.3	98.3	72.0	95.6
ST	95.1	70.3	98.0	72.6	**96.7**	93.8	56.5	98.1	62.7	92.4	95.3	71.4	98.0	73.3	95.2
	94.8	—	—	—	—	—	—	—	—	—	*95.0*	*77.2*	*97.0*	*73.4*	*95.6*

	Bankhead method					*Nguyen method*					Second manual segm.			
—	Acc	Sn	Sp	MCC	AUC	Acc	Sn	Sp	MCC	AUC	Acc	Sn	Sp	MCC
AR	93.8	56.9	97.6	50.0	85.7	93.8	52.4	98.1	58.7	87.5	93.0	57.6	96.7	57.0
CH	94.0	64.4	97.4	65.4	91.7	94.4	66.5	97.5	67.4	93.5	95.6	77.0	97.8	75.7
DR	94.0	63.1	98.6	70.7	90.5	94.5	67.8	98.4	73.3	93.4	94.7	77.6	97.3	76.0
	93.7	*70.2*	*97.2*	—	—	*94.1*	—	—	—	—	—	—	—	—
HR	95.6	71.2	98.1	71.8	92.1	**95.8**	72.0	98.2	73.3	94.7	—	—	—	—
ST	95.2	69.2	98.2	72.3	93.7	**95.5**	71.5	98.3	74.3	95.8	93.5	89.5	93.8	72.2
	—	—	—	—	—	*93.2*	—	—	—	—	—	—	—	—

AUC optimized

	Soares method					*Sofka method*					*Azzopardi method*				
—	Acc	Sn	Sp	MCC	AUC	Acc	Sn	Sp	MCC	AUC	Acc	Sn	Sp	MCC	AUC
AR	93.5	52.6	97.7	57.1	90.9	93.3	44.4	98.3	53.3	86.3	**93.8**	54.5	97.8	59.2	**92.4**
CH	**94.5**	69.7	97.3	68.9	**96.4**	93.2	50.9	98.0	58.0	90.8	94.2	67.0	97.2	66.7	95.9
DR	**94.7**	72.0	98.0	74.9	96.2	93.5	60.9	98.2	67.9	91.4	94.4	70.1	98.0	73.4	**96.4**
HR	**95.7**	73.0	98.1	73.1	**97.1**	94.3	58.3	97.8	61.4	93.7	95.4	68.1	98.1	70.4	96.7
ST	95.1	71.3	97.9	72.7	96.8	93.8	56.4	98.1	62.7	92.4	95.0	68.7	98.0	71.5	**97.0**

	Bankhead method					*Nguyen method*				
—	Acc	Sn	Sp	MCC	AUC	Acc	Sn	Sp	MCC	AUC
AR	92.9	47.3	97.7	52.9	88.8	93.7	53.1	97.9	58.2	89.5
CH	94.0	64.4	97.4	65.4	91.7	94.0	64.3	97.4	65.2	94.7
DR	93.1	59.4	98.1	66.0	91.1	94.4	67.2	98.4	73.0	94.2
HR	95.4	71.3	97.9	71.0	94.1	95.5	68.8	98.3	71.2	95.8
ST	94.7	66.0	98.1	69.5	95.0	**95.3**	71.4	98.1	73.5	96.3

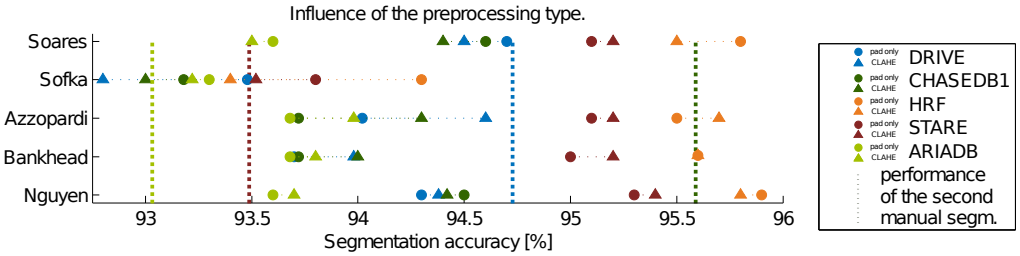


Figure 3.2: A comparison of the segmentation performance: pad only (circles) and CLAHE (triangles). The databases are marked with different colours. Solid dots and triangles mark the algorithm accuracy. Vertical lines mark the performance of the second observer when available.

images for the *Nguyen, Azzopardi and Bankhead methods*, and the results based on pad only for the *Soares and Sofka methods*.

A comparison of the methods reveals relatively similar performance. For the four methods (except the *Sofka method*), the absolute difference between the best and the worst accuracy on individual databases was up to 0.5 in percentage units. No method seemed to be clearly superior. The best and second best performance were achieved by the *Azzopardi and Bankhead methods* on ARIADB, by the *Soares and Nguyen methods* on CHASEDB1, by the *Soares and Azzopardi methods* on DRIVE, by the *Nguyen and Soares methods* on HRF and by the *Nguyen and Azzopardi methods* on STARE.

The search for optimal parameters brought about a small improvement in the performance of the algorithms compared to the performance published in the original papers. Nguyen et al. [42] obtained lower τ value than was obtained in the experiments presented in this chapter which led to significantly worse performance on STARE. Compared to the original papers, the performance of the *Bankhead and Nguyen methods* were slightly improved by CLAHE preprocessing.

Comparison with state of the art

Here we provide a brief comparison of the tested methods with state-of-the-art methods which are not available with implementation. Papers for the comparison that were published before year 2011 were gathered from the review by Fraz et al. [8]. The more recent papers were gathered from the list of papers that cite [64], [65, 66, 67, 68] – the publications introducing the databases reviewed in Section 2.6. It was observed that many methods report high performance but without providing a clear methodology for performance assessment. To ensure that the comparison is fair, only methods explicitly stating that performance was measured on pixels inside the FOV were included.

Typically the performance of retinal vessel segmentation algorithms is reported on DRIVE and STARE and, thus, many results are available using those databases. Few methods were identified that also reported performance on CHASEDB1 and HRF. A comparison of those state-of-the-art methods for which accuracy was reported is presented in Tables 3.5, 3.6 and 3.7. When sensitivity and specificity were also provided, the per-

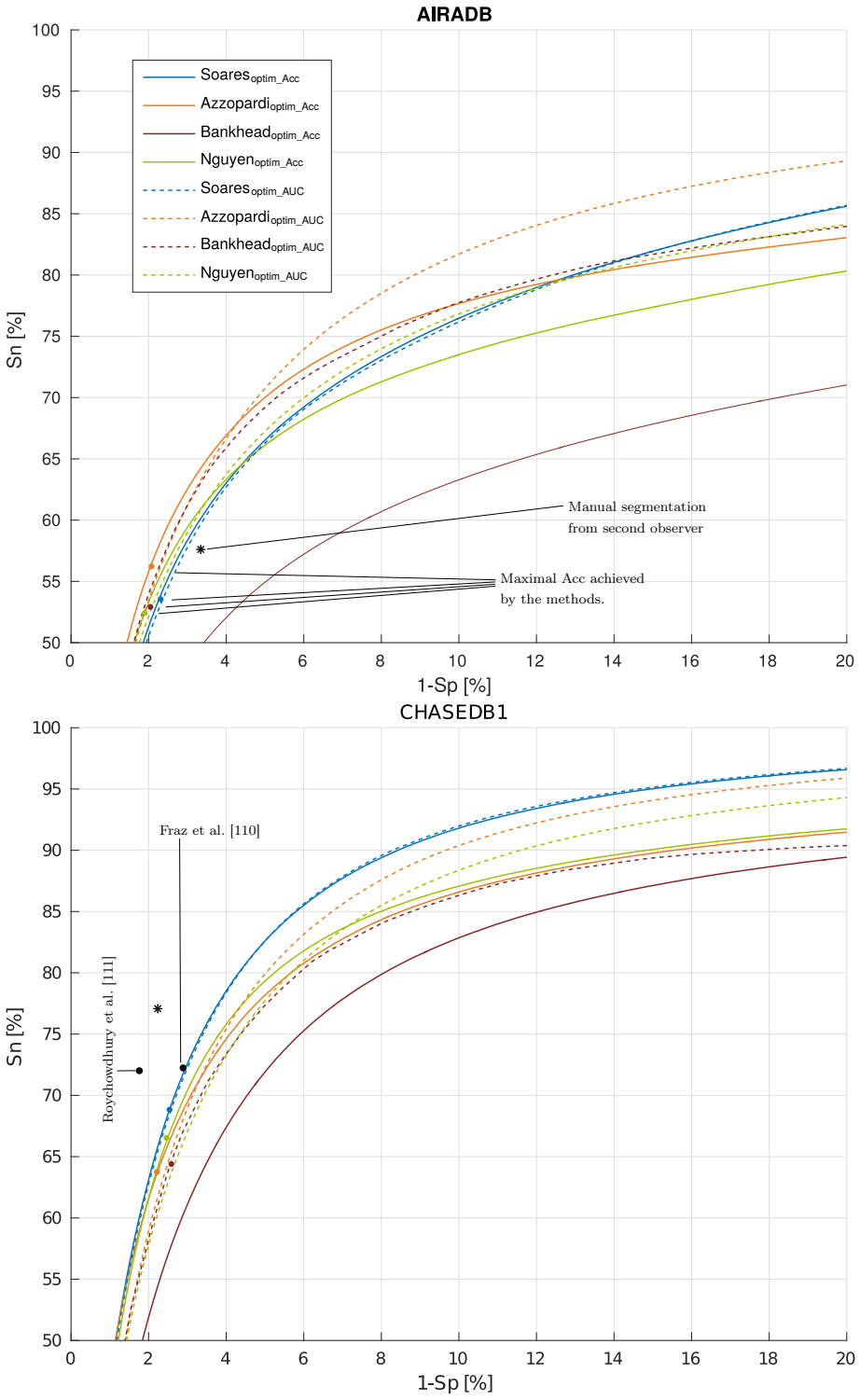


Figure 3.3: The ROC characteristics of the studied methods. Manual segmentation by second observer is marked with an asterisk. The ROC curves correspond to the parameters optimized by Acc (the solid line) and AUC (the dotted line). The *Bankhead method* is different due to its post-processing: one is the ROC curve of the IUWT response (the solid line) and the other is the convex hull of all possible performances from the parameter search (the dotted line).

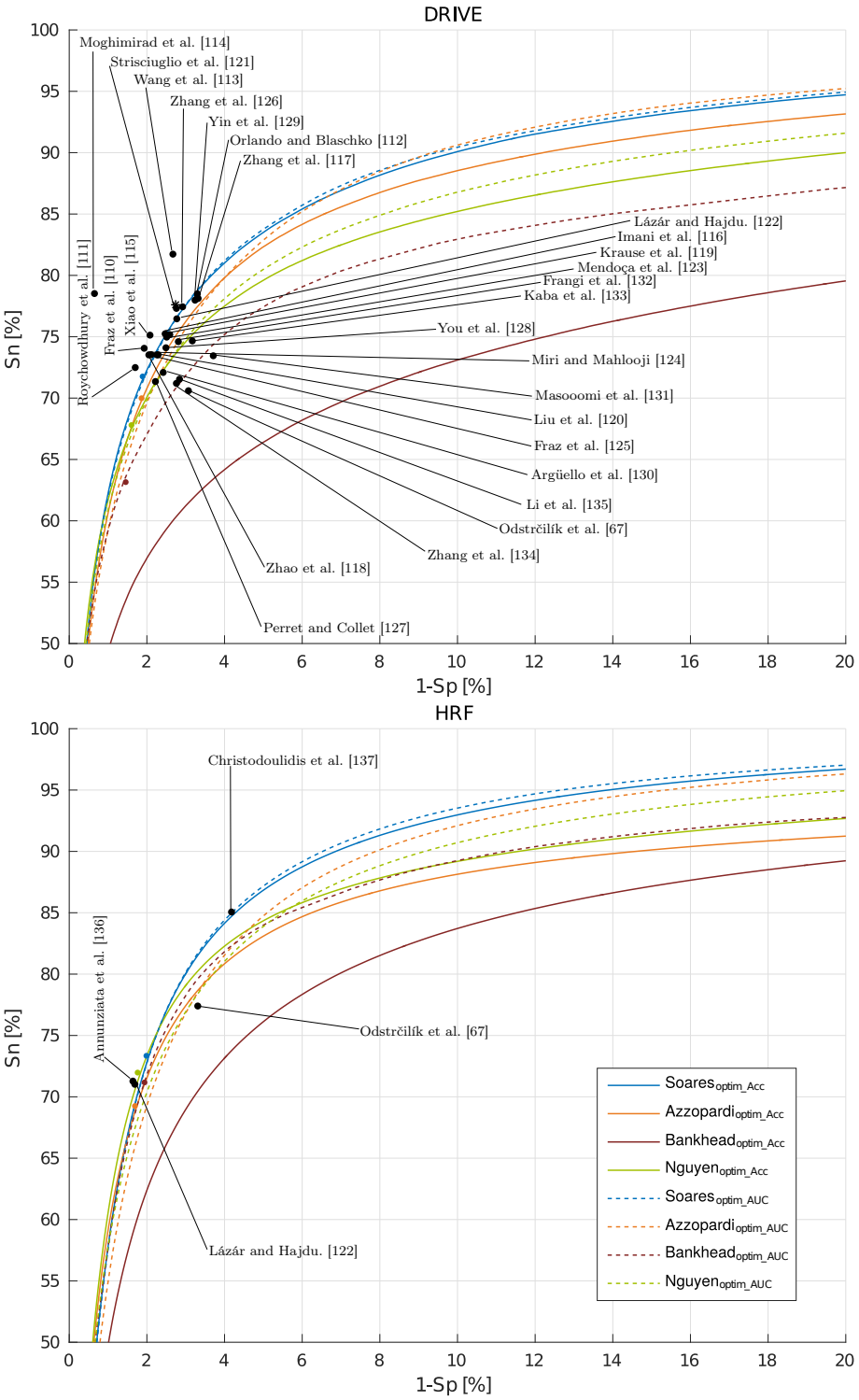


Figure 3.4: The ROC characteristics of the studied methods. Manual segmentation by second observer is marked with an asterisk. The ROC curves correspond to the parameters optimized by Acc (the solid line) and AUC (the dotted line). The *Bankhead method* is different due to its post-processing: one is the ROC curve of the IUWT response (the solid line) and the other is the convex hull of all possible performances from the parameter search (the dotted line).

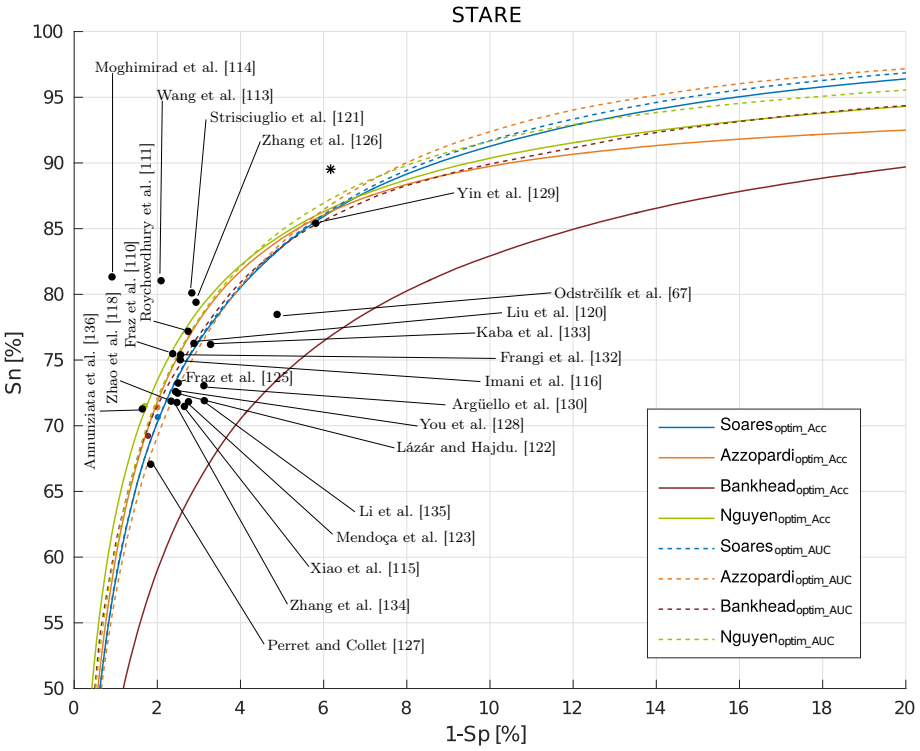


Figure 3.5: The ROC characteristics of the studied methods. Manual segmentation by second observer is marked with an asterisk. The ROC curves correspond to the parameters optimized by Acc (the solid line) and AUC (the dotted line). The *Bankhead method* is different due to its post-processing: one is the ROC curve of the IUWT response (the solid line) and the other is the convex hull of all possible performances from the parameter search (the dotted line).

formance was plotted in Figures 3.3 – 3.5. The latter way of comparing the methods enables a clearer and fairer way of comparison.

Table 3.5: An overview of state-of-the-art methods evaluated on DRIVE and STARE. The methods are sorted by mean performance on both databases.

	DRIVE				STARE			
Algorithm	Sn	Sp	Acc	AUC	Sn	Sp	Acc	AUC
Wang et al. [113]	81.7	97.3	97.7	94.8	81.0	97.9	98.1	97.5
Moghimirad et al. [114]	78.5	99.4	96.6	95.8	81.3	99.1	97.6	96.8
Imani et al. [116]	75.2	97.5	95.2	—	75.0	97.5	95.9	—
Al-Rawi et al. [138]	—	—	95.4	94.4	—	—	—	—
Lam et al. [139]	—	—	94.7	96.1	—	—	95.7	97.4
Liu et al. [120]	73.5	97.7	94.7	—	76.3	97.1	95.7	—
Annunziata et al. [136]	—	—	—	—	71.3	98.4	95.6	96.6
Roychowdhury et al. [111]	72.5	98.3	95.2	96.2	77.2	97.3	95.2	96.9
Fraz et al. [110]	74.1	98.1	94.8	97.5	75.5	97.6	95.3	97.7
Xiao et al. [115]	75.1	97.9	95.3	—	71.5	97.4	94.8	—
Zhang et al. [117]	78.1	96.7	95.0	—	—	—	—	—
Strisciuglio et al. [121]	77.3	97.2	94.7	95.9	80.1	97.2	95.4	96.3
Zhao et al. [118]	73.5	97.9	94.8	—	71.9	97.7	95.1	—
Krause et al. [119]	75.2	97.4	94.7	—	—	—	—	—
Soares method	71.7	98.1	94.7	96.1	70.3	98.0	95.1	96.7
Zhang et al. [126]	77.4	97.1	94.5	—	79.4	97.1	95.1	—
Nguyen method	67.8	98.4	94.5	93.4	71.5	98.3	95.5	95.8
Azzopardi method	70.0	98.1	94.5	95.6	71.4	98.0	95.3	95.2
Orlando and Blaschko [112]	78.5	96.7	—	—	—	—	—	—
Staal et al. [66]	—	—	94.4	95.2	—	—	95.2	96.1
Miri and Mahlooji [124]	73.5	98.0	94.6		—	—	—	—
Fraz et al. [125]	73.5	97.7	94.5	96.7	73.3	97.5	95.0	96.7
Perret and Collet [127]	71.4	97.8	94.4	—	67.1	98.2	95.1	—
Lázár and Hajdu. [122]	76.5	97.2	94.6	—	72.5	97.5	94.9	—
You et al. [128]	74.1	97.5	94.3	—	72.6	97.6	95.0	—
Tagore et al. [140]	—	94.2	95.3	—	—	—	95.0	96.1
Masooomi et al. [131]	73.5	96.3	94.3	—	—	—	—	—
Frangi et al. [132]	74.6	97.2	94.2	—	75.4	97.4	95.0	—
Mendoça et al. [123]	75.0	97.5	94.6	—	71.8	97.3	94.6	—
Bankhead method	63.1	98.6	94.0	90.5	69.2	98.2	95.2	93.7
Argüello et al. [130]	72.1	97.6	94.3	—	73.1	96.9	94.5	—
Zhang et al. [134]	71.2	97.2	93.8	—	71.8	97.5	94.8	—
Kaba et al. [133]	74.7	96.8	94.1	—	76.2	96.7	94.6	—
Yin et al. [129]	78.0	96.8	94.3	—	85.4	94.2	93.3	—
Sofka method	60.9	98.2	93.5	91.5	56.5	98.1	92.4	93.8
Li et al. [135]	71.5	97.2	93.4	—	71.9	96.9	94.1	—
Odstrčilík et al. [67]	70.6	96.9	93.4	95.2	78.5	95.1	93.4	95.7

Table 3.6: An overview of state-of-the-art methods evaluated on CHASEDB1.

Algorithm	Sn	Sp	Acc	AUC
Roychowdhury et al. [111]	72.0	98.2	95.3	95.3
Fraz et al. [110]	72.2	97.1	94.7	97.1
Soares method	69.0	97.7	94.6	96.4
Nguyen method	66.5	97.5	94.4	93.5
Azzopardi method	63.7	97.8	94.3	93.2
Bankhead method	64.4	97.4	94.0	91.7
Sofka method	45.6	98.3	93.0	89.1

Table 3.7: An overview of state-of-the-art methods evaluated on HRF.

Algorithm	Sn	Sp	Acc	AUC
Cheng et al. [141]	70.4	98.6	96.1	—
Soares method	73.4	98.0	95.8	97.0
Christodoulidis et al. [137]	85.1	95.8	94.8	—
Nguyen method	72.0	98.2	95.8	94.7
Annunziata et al. [136]	71.3	98.4	95.8	—
Azzopardi method	69.3	98.3	95.7	95.6
Bankhead method	71.2	98.1	95.6	91.3
Lázár and Hajdu. [122]	71.0	98.3	95.3	—
Odstrčilík et al. [67]	77.4	96.7	94.9	96.7
Sofka method	58.3	97.8	94.3	93.7

3.4.2 Parameter prediction

Parameters obtained from the state-space search are reported in Table 3.8. These parameters correspond to the results in Table 3.4.

In order to compare the resolution-related models, three best models based on the range of R^2 were selected from groups of single-, two- and three-predictor models. The performance of the models is summarized in Table 3.9. It was concluded from the results that similar groups of models yielded high results in both cases of parameter optimization by Acc and AUC, which strengthens confidence in the generalizability of the models. For single predictor models, d_r and d_{v1} were the most informative predictors with median $R^2 > 0.95$ ranging from 0.39–0.99. For models with two predictors the best combinations were $\{d_{v1}, d_{od}\}$ and $\{d_r, d_{v1}\}$ with median $R^2 > 0.96$ and the range raised to 0.6–0.9. For models with three predictors the parameters were $\{d_{v1}, d_{v2}, d_{od}\}$ with median $R^2 > 0.97$ and the range 0.94–1. In summary – all the methods scale well among images of different resolutions. The parameters r_1, r_2 of the *Azzopardi method* were the most difficult to predict. The parameters optimized according to AUC were easier to predict than parameters optimized by Acc (the lower range of R^2).

The parameters from the non-resolution-related group were more difficult to model by the chosen predictors. Tables with the corresponding assessment were placed in Appendix I. Parameter-wise, the *Azzopardi method*'s σ_1 and σ_2 optimized by Acc had low R^2 even for models with three predictors (maximally 0.87). When optimized by AUC

Table 3.8: Parameters corresponding to the results reported in Table 3.4. Note that all the τ within the AUC optimized parameters were optimized according to Acc. The parameters which were acquired as a two sets (CHASEDB1, STARE) were averaged. Optimization of the parameters by ∼AUC is not applicable to the *Bankhead method*.

Acc optimized	ARIADB	CHASEDB1	DRIVE	HRF	STARE
Soares					
Λ_{mor}	{2,3,5}	{3,7,9}	{2,4}	{5,13,15}	{2,4,5}
n_s	2e5	3e5	2e5	3e5	2e5
Sofka					
τ	1.86	1.3	0.82	3.74	1.75
Azzopardi					
σ_1, σ_2	3.3, 1.6	5.3, 3.9	2, 1.6	7.2, 6.4	2.8, 1.6
r_1, r_2	19, 27	21, 22	8, 25	23, 44	12, 28
σ_{01}, σ_{02}	1, 0.5	2.5, 0	3, 1.5	0.5, 0	0.5, 0.5
a_1, a_2	0.4, 0.1	0.2, 0	0.5, 0.1	0.4, 0	0.5, 0.1
τ	0.14	0.16	0.16	0.16	0.15
Bankhead					
Λ_{ban}	{2,3}	{3,4}	{2,3}	{3,4}	{2,3}
p_t	0.12	0.12	0.12	0.12	0.12
ξ_s	522	780	150	1030	270
Nguyen					
W	19	33	17	45	17
ω	18	16	4	22	16
τ	1.05	0.95	0.9	1.05	1.05

AUC optimized	ARIADB	CHASEDB1	DRIVE	HRF	STARE
Soares					
Λ_{mor}	{2,3,6}	{3,6,10}	{2,3,5}	{5,10,16}	{2,4,7}
n_s	2e5	3e5	2e5	3e5	2e5
Sofka					
τ	1.86	1.3	0.82	3.74	1.75
Azzopardi					
σ_1, σ_2	4.9, 2.4	7.1, 2.8	2.8, 1.8	7.4, 6.8	2.7, 2.8
r_1, r_2	16, 24	21, 28	8, 16	26, 50	12, 24
σ_{01}, σ_{02}	4.5, 1	9, 1.3	4.5, 3	9.5, 0	1.3, 3
a_1, a_2	0.4, 0	0, 0.1	0.1, 0.1	0.4, 0.1	0.2, 0.1
τ	0.18	0.185	0.18	0.21	0.2
Bankhead	—	—	—	—	—
Nguyen					
W	20	39	19	59	19
ω	2	2	2	2	6
τ	1.25	1.15	0.95	1.25	1.1

Table 3.9: The linear models of the resolution-related parameters based on different predictors. Assessment of the models was done by averaging R^2 over the parameters. The models marked with an asterisk were statistically significant for all the parameters.

Predictors (Acc)	R^2 median (min, max)	Predictors (AUC)	R^2 median (min, max)
d_r	0.91 (0.41, 0.99)	d_{v1}*	0.91 (0.72, 0.99)
d_{v1}	0.94 (0.39, 0.99)	d_{v3}	0.83 (0.58, 0.96)
d_{od}	0.86 (0.35, 0.99)	d_r	0.95 (0.54, 0.99)
d_{v1}, d_{od}	0.98 (0.69, 1)	d_{v1}, d_{od}*	0.95 (0.91, 1)
d_r, d_{v1}	0.96 (0.60, 1)	d_r, d_{v1}	0.96 (0.90, 1)
d_{df}, d_{nf}	0.88 (0.50, 0.97)	d_{v1}, d_{nf}	0.95 (0.82, 1)
d_{v2}, d_{df}, d_{nf}	1.0 (0.95, 1)	d_{v1}, d_{v2}, d_{od}	0.97 (0.94, 1)
d_{v1}, d_{v2}, d_{od}	0.99 (0.94, 1)	d_n, d_{v1}, d_{od}	0.99 (0.90, 1)
d_{v1}, d_{df}, d_{nf}	0.98 (0.94, 1)	d_{v1}, d_{v3}, d_{od}	0.97 (0.88, 1)

the predictability was better for the models with one or two predictors. The parameters a_1 and a_2 were predictable for those optimized by Acc (R^2 0.65–0.99) but were not predictable when AUC was optimized. The parameter ω of the *Nguyen method* was also not predictable by models with less than three predictors in the case of using both approaches to optimization.

The parameters p_t, Λ_{ban} of the *Bankhead method* spans a small range of values and the fixed value of 0.12 was used as the modelled value. The assessment of a selected model using classification accuracy can be found in Table I.6. The practical application of the proposed models is described in Chapter 3.

3.5 Discussion

As a consequence of lacking a standard evaluation methodology for automatic retinal image processing methods [142], factors affecting the outcome of the evaluation (for example, test set size, image preprocessing, evaluation metrics) vary among studies. Here, segmentation accuracy was used as the primary measure for comparing segmentation performance. It is the most widely used measure for assessing the segmentation performance and it makes interpreting the results intuitive. Using accuracy alone can lead to simplifying conclusions when, in general, the accuracy of the segmentations with different sensitivity are compared. Typically, human raters tend to produce results with higher sensitivity than automatic algorithms, but their accuracy can be lower. Performance on the STARE database is a typical example [37]. Automatic methods, on the other hand, might not be able to compete with manual segmentation at the same sensitivity level. Therefore, we found it necessary to also provide the ROC characteristics (Figures 3.3 – 3.5) of the algorithms to demonstrate the relative performance of the methods at different sensitivity.

The performance of all the methods, except the *Sofka method*, was very much similar on all the databases studied, and the order of best-performance varied based on the different databases analyzed. The *Soares method* provides a slightly better AUC – it offers better accuracy when higher sensitivity is considered. The comparison with the state-of-the-art methods (Figures 3.3 – 3.5) revealed only few significantly better methods and showed the methods to perform closely to each other. There might thus be a need to improve the whole assessment methodology. One option could be to use skeletonized vessels [38], which results in higher demands on the detection of narrower vessels.

As stated above, ROC curves offer a meaningful way to compare the performance of the automatic methods and manual segmentations. It was shown in Figures 3.3 – 3.5 that the manual segmentation of CHASEDB1 and STARE outperform the automatic segmentation markedly, and they can be regarded as challenging datasets. For DRIVE, the Soares and Azzopardi methods offer performance very close to the manual segmentation, and for ARIA, manual segmentation is outperformed by all the methods.

Considering the advantages and disadvantages of the methods reviewed from the original publications in Subsection 2.3, most authors claimed fast segmentation. The exact segmentation times were not measured in our study, but it is possible to conclude that each method classified an image of STARE or DRIVE within the order of seconds, the fastest method being the *Bankhead method*, followed by the *Nguyen and Azzopardi methods*. It is worth noting that the training phase of the *Soares method* was speed up approximately 20 times by optimizing the source code.

The *Sofka method* was the only method designed to prevent false positive classifications around high-contrast areas like pathologies or the OD. The performance of the method was, however, well below the others. Qualitative inspection of the results showed that the reason for the poor performance may mainly be because of the method is not well suited for pixel-wise classification [38]. The resulting vesselness gains high values near the vessels thus producing classification wider than that one used as the GT. Therefore, in [38] the method was used to detect the vessel centrelines. However, the method was not inspected closer as it is provided as an executable binary, and our focus was more on the methods provided with the source code.

Qualitative inspection[2] of the other methods showed that with the *Soares method* false positive detections at the edge of the OD are rare but appear more often with the unsupervised methods. This was, however, largely corrected by applying CLAHE preprocessing. The preprocessing, on the other hand, increased the noise in the images as well as the subsequent classifications. The *Azzopardi method* was the most resistant to the noise and the only one benefiting significantly from the preprocessing. Uneven illumination does not seem to influence the classification of the unsupervised methods. A more important factor appears to be that the vessels in the over- or under-illuminated areas have reduced contrast. On the other hand, the *Soares method* is affected more by the uneven illumination, possibly due to fact that there are not enough examples of such images in the training data.

Considering the individual methods, the *Soares method* provides very good performance if part of the classified data is available for training. Training on a different database leads to a performance drop [37] and when databases with different image resolution are

[2]The vessel segmentation in all the tested images is available at `http://www.it.lut.fi/mvpr/medimg`.

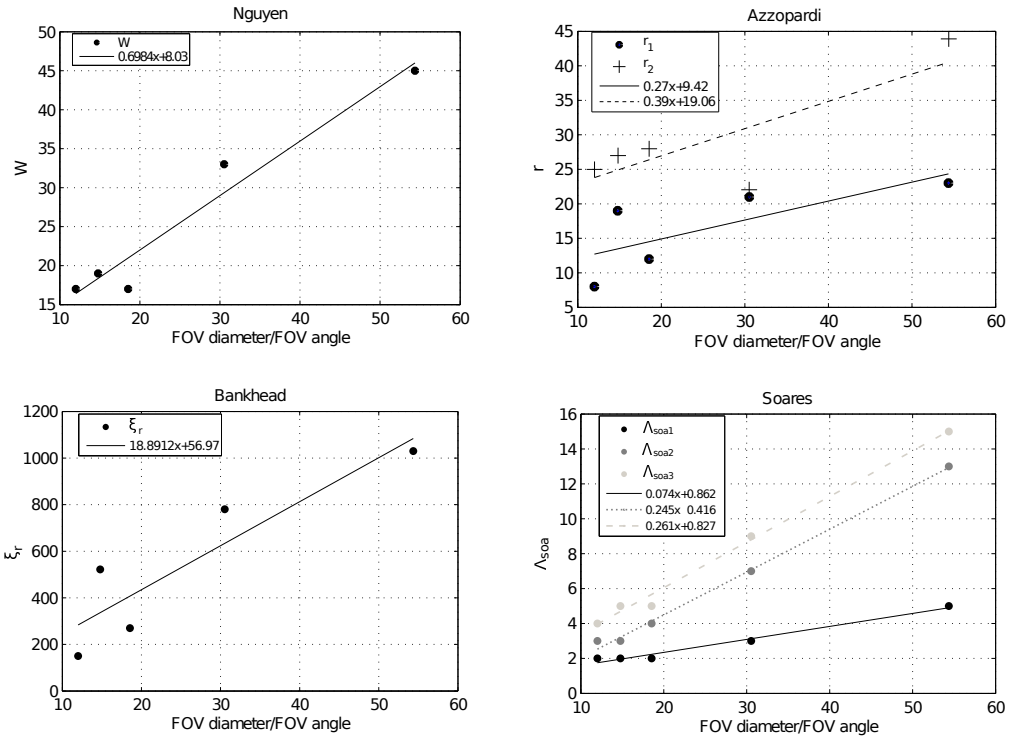


Figure 3.6: An Illustration of the linear models. Selected parameters are correlated with the angular resolution of the databases.

considered, it is infeasible to properly train the classifier unless the images are resampled. The *Soares method*, when properly trained, provides high AUC which is valuable when segmentation with higher sensitivity is desired. Of the unsupervised methods, the *Azzopardi method* yielded the highest AUC, which in some databases was even comparable with the *Soares method*. The *Azzopardi method* seems to produce robust segmentation of the wider vessels and also has the best robustness against pathologies and other high-contrast areas. Narrow vessels cause the lower response of the filter and are usually missed. The *Nguyen method* provides performance similar to the *Azzopardi method*, but loses performance with pathologies with high contrast. On the other hand, it has a higher response to narrow vessels thus providing a more balanced segmentation. The AUC of the *Nguyen method* is lower than that of the *Soares and Azzopardi methods*. Lastly, the *Bankhead method* provides the worst performance although not by a high margin. At the same time, it is the fastest method. Also the *Bankhead method* exhibits performance improvement from simple post-processing, such as the removal of isolated objects up to a certain size, as can be seen in the performance gain of the method in Figures 3.3 – 3.5.

Most of the parameters of the methods proved to be possible to model and predict by using the database resolution and the expected percentage of vessel pixels in the GT. Also, the strong correlation of most of the parameters with the predictors (Figure 3.6) supports this hypothesis. As a result, the proposed models are expected to provide

close-to-optimal parameters for the methods. However, the models were not validated on databases not used to establish the models. This was due to the limited number of the databases. Therefore, the models could fail on other, unseen databases. In that case, the *Bankhead method* seems to be the option with the smallest number of choices required for parameter settings and the best robustness against parameter variation as it needs a small range of wavelet coefficients, uses percentile-based thresholding and the size of the removed objects can be defined proportionally to the size of the ROI.

With regard to the sets of the tested parameter values, the range of limits were carefully selected. In the case of the *Bankhead and Nguyen methods*, the parameters are expected to be sampled properly in order to get close-to-optimal settings on each database. The Azzopardi method is harder to optimize because it has eight parameters. The whole parameter subspace defined by the ranges was thus difficult to evaluate, so there might be room for improvement. Azzopardi et al. [39] have, however, already optimized the parameters for DRIVE, STARE and CHASEDB1. With the *Soares method*, no parameter search was made for the parameters of the Morlet wavelet other than Λ_{mor}, which was limited to integer values. This provides room for further improvement of the Morlet wavelet-based vessel segmentation methods and the current tested method in particular; setting the ϵ and k_0 parameters can produce a filter with similar properties as the line operator that is implemented in the *Nguyen method*. The advantages of this type of response could be integrated into the framework of the *Soares method*. One observation with respect to the wavelet filter mask is that between the Λ_{mor} values 1 and 1.5, the wavelet mask has a different mean value for different orientations and using a different normalization of the mask might enable use of the response with a wavelet level size <2.

Chapter IV

Retinal vessel quantification

4.1 Introduction

Retinal vessel features like diameter, tortuosity and fractal dimension have been identified to be connected with various retinal and systemic diseases and proposed for the early identification of these diseases [26]. The employment of the computerized systems to automate the manual processing of the images has, thus, been of high interest as in the case of vessel segmentation a large body of research has been oriented towards methods for the automatic assessment of the vessel features.

In this chapter, an approach to automation of the AVR measurement and its assessment on a new retinal database is proposed. The approach consists of the application of the vessel segmentation methods to the database, separation of the vessels into individual segments, generation of AV GT and, finally, separating the vessels into arteries and veins with subsequent selection of the final vessels and automatic estimation of the AVR. The ROI for the estimation of AVR is placed in the usual location, inside a circular ring within 1–1.5 ODDs from the OD centre. Locating the centre and measuring the diameter of the OD were manual in this work.

In Chapter 2, the clinical background of the retinal vessel quantification was described, with a particular focus on the measurement and applications of AVR: A review of the vessel features – or retinal characterization measures – was provided in Section 2.2. Then in Section 2.4 we provided a review of the automatic methods for the AV classification of retinal vessels. In addition, Section 2.5 provided an overview of the automated frameworks for AVR estimation.

In this chapter, the database used in this study will be introduced in Section 4.2. Throughout Section 4.3, the parts of the AVR framework are introduced: first the prediction models of the vessels segmentation parameters are applied to obtain the vessel segmentation on our database, next the approach to the vessel tracing and GT estimation is described, then the section focuses on our approach to the automatic classification of the segmented vessels into arteries and veins, and finally the approach to estimate the

final AVR is covered. The results of the AV classification using our selection of features and the comparison of the automatically and manually computed AVRs against the blood pressure is presented in Section 4.4. The results are discussed in Section 4.5.

4.2 Data

4.2.1 The Savitaipale database

The database used in this work was based on the Savitaipale study – a population-based follow-up cohort study on the risk factors of type 2 diabetes and screening for abnormal glucose metabolism in primary health care. The study consisted of adults born 1933–1956 and residing in the Savitaipale municipality in Eastern Finland. The participants of the study were examined at baseline during 1996–1999 and with follow-up during 2006–2008. The retinal photographs were taken during the follow-up. In total 1561 participants were invited to the study, off whom 1168 agreed to participate.

Images in the database are RGB colour fundus photographs of the size 2544x1696 pixels, stored in JPEG (with optically visible quality degradation by compression) format with 8-bit colour depth per colour channel. Four retinal photographs are available per person – one macula and one OD centred picture from both eyes. From the clinical parameters measured during the study, systolic and diastolic blood pressure, age, BMI and sex were made available for use in the presented work. A summary of the clinical data is provided in Table 4.1.

Manual measurement of the AVR, made by a person trained in the task, was provided with the database. It was made on 915 images which had sufficient quality. When doing the measurements, the grader was masked to the subjects' characteristics and endpoints. Images of the right eye were preferably measured; when an image was not of sufficient quality then the image of left eye was used. The diameters of the six largest arterioles and six largest venules (in pixels) were measured with a computer program (Adobe Photoshop CS 11) in an area of 0.5 to 1 standard ODD from the OD edge. Standard ODD in the study was 301.7 pixels, which was established based on a random sub-sample of 100 participants. Typically the branches were measured in order to fulfil the preferred number of the vessels (six).

4.3 Methods

Throughout this section, the application of the vessel segmentation prediction, as proposed in Subsection 3.3.2, to our database is described in Subsection 4.3.1. Approach used to trace the edges of the segmented vessels and measure the vessel width is summarized in the Subsection 4.3.2. Approach to estimation of the AV ground truth for the segmented vessels is summarized in the Subsection 4.3.3. The most important parts of the AVR estimation framework are the approach to the vessel classification into arteries and veins which is covered in the Subsection 4.3.4, and the approach to the final AVR estimation, which is covered in the Subsection 4.3.5.

Table 4.1: A summary of the clinical measurements of the participants in the Savitaipale study and manual measurement of the AVR. The values are reported as the mean and SD values.

	Men	Women
No. of images	419	466
Age	62.5 (6.4)	63 (6.7)
BMI	27.4 (4.4)	27.4 (4.9)
Systolic pressure [mmHg]	133.3 (15.0)	132.8 (14.6)
Diastolic pressure [mmHg]	83.0 (8.5)	80.3 (8.4)
AVR	0.626 (0.06)	0.643 (0.05)

4.3.1 Prediction of the vessel segmentation parameters

No manual segmentation of the retinal vessels was available for the Savitaipale database and, thus, methods for the vessel segmentation and prediction of their parameters from Subsection 3.3.2 were applied. In order to obtain reasonable vessel segmentation, several combinations of the segmentation methods and the prediction models of their parameters were tested. The *Azzopardi* and *Nguyen methods* were selected for testing because they offer slightly better performance than the *Bankhead method* and no training examples were available to train for the *Soares method*.

The selection of the model for the prediction of the method's parameters and the assessment of it was done by subjective qualitative comparison of the segmentation obtained using a subset of the models described in Subsection 3.3.2. Models based on the following sets of predictors were tested:

- $\{d_{v1}\}$, referred to as AUC_1 or Acc_1
- $\{d_{v1}, d_{od}\}$, referred to as AUC_2 or Acc_2
- $\{d_{v1}, d_{v2}, d_{od}\}$, referred to as AUC_3 or Acc_3

AUC/Acc marks the measure used for optimizing the predicted parameters. Those were the best among the models for predicting the resolution-related parameters. Although these models are confirmed to only work well for the resolution-related parameters they were used for predicting all the parameters because no important difference was observed when a different model was setup for prediction of the non-resolution-related parameters.

Five images were randomly selected from the Savitaipale database for comparison of the models, segmented and compared by the author. Illustrations of the differences in the obtained segmentations can be found in Figure 4.1. The *Azzopardi method* was selected as the segmentation algorithm because it appeared to better separate close vessels and better delineate the edges of the vessels in general, as illustrated in Figure 4.1 at the bottom. Comparison between the models 1, 2 and 3 did not reveal any big differences in the segmentation, although the model $\{d_{v1}, d_{v2}, d_{od}\}$ appeared to give slightly more robust results in edge detection in the presence of noise and central vessel reflex (an

illustration is given in Figure 4.1, on the top row). A decision between using the AUC-based model or the Acc-based model led to the selection of the AUC-based one because of its higher robustness against noise and better detection in the presence of central vessel reflex (an illustration of this can be found in Figure 4.1, on the middle row).

The *Azzopardi method* with the prediction model AUC_3 was used for the final classification. Furthermore the post-processing method proposed by Bankhead et al. [40] (the removal of connected components up to a certain size and the filling of holes up to a certain size) was applied after the segmentation. The threshold to obtain binary classification from the *Azzopardi method* was determined manually as well as the settings of the post-processing: $\tau = 0.95$, $\xi_s = 2000$, $\xi_h = 200$.

4.3.2 Vessel tracing

Tracing of the vessel edges and measurement of their diameter was done using the method proposed by Bankhead et al. [40], employing their Matlab implementation. In the method, the traced vessel segments are derived from the binary vessel segmentation which is thinned with subsequent detection of junction points that serve to delineate individual segments. The thinned segment centreline pixels are then interpolated using splines. The spline is used to establish interpolated straightened vessel picture. This picture is blurred with the directional Gaussian filter – prolonged in the direction of the vessel. The edges of the vessel are then searched from the centreline in a perpendicular direction; the edge is established as the point with maximal gradient around the centreline.

4.3.3 Arterio-venous ground truth estimation

In order to train and validate our approach to the AV classification of the vessel segments the GT measurement of the vessel classes was needed. Having the vessels segmented and separated into segments, as described in Sections 4.3.1 and 4.3.2, a tool was implemented that allowed the labelling the segments as *artery*, *vein*, *both*, *unknown* and *no-vessel*. At the same time, it is possible to label endpoints of each segment as *head* or *tail* (or *head-head*, *tail-tail* and *counter-current* for some rare types of vessel orientation).

The location and diameter of the OD was also provided manually in our experiments and another tool was implemented for that. Each OD was delineated by set of points that were fitted by an ellipse. The average length of the axes of the ellipse were used as an estimation of half an ODD.

4.3.4 Classification into arteries and veins

This section describes the approach to AV classification in the presented framework. Supervised classification was chosen for the task with the LS-SVM [47] classifier previously proposed by Relan et al. [44]. The supervised classification was chosen because it generally produces better and more robust results when appropriate training data are available, which was also the case in our experiments. Classification features were estimated from averaged vessel profiles where the averaged profile was computed as the mean of p_n individual profiles (such a sub-profile can be seen in Figure 4.2, delineated

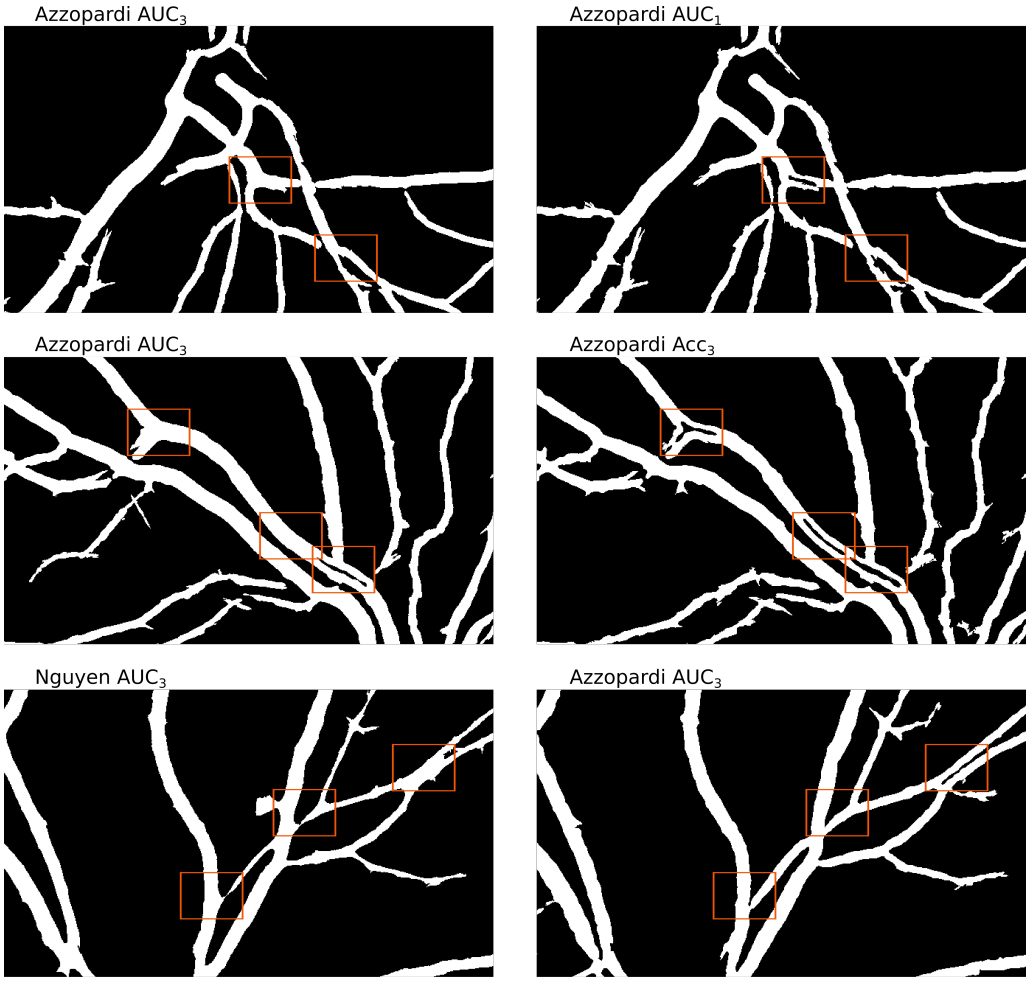


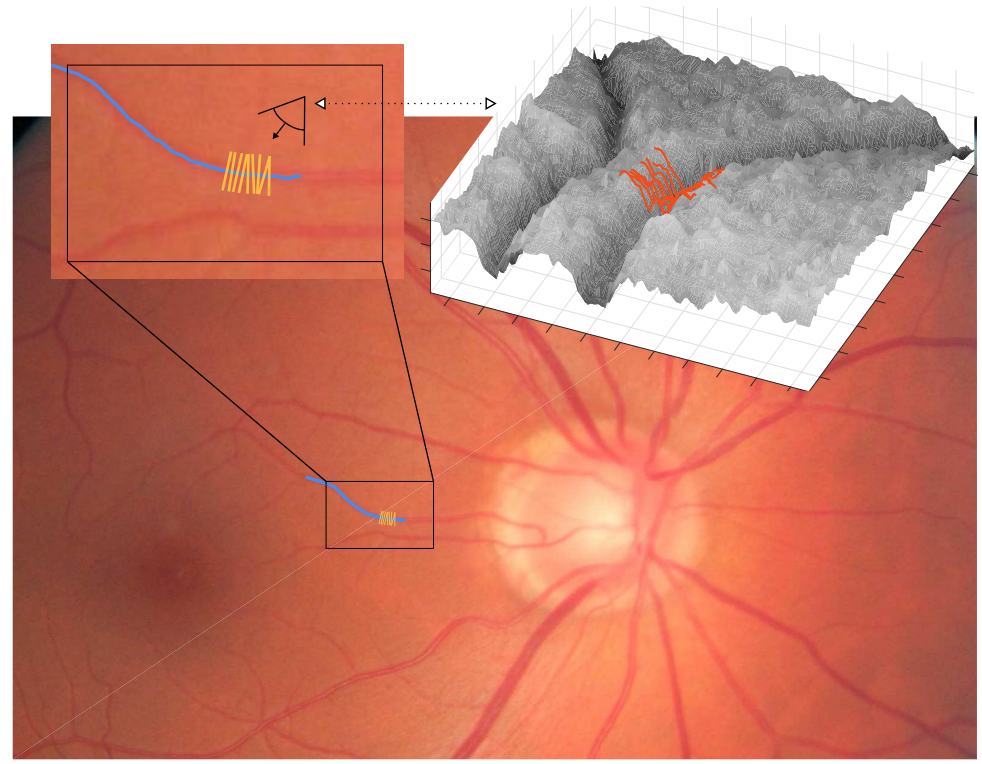
Figure 4.1: An illustration of the properties of the distinct prediction models for the vessel segmentation that were considered for use on the Savitaipale database. Areas with the important properties of the models are delineated by coloured rectangles. The depicted properties are: on the top row – more robust edge detection and better robustness in the presence of the central vessel reflex of the model AUC_3; on the middle row – better robustness in the presence of central vessel reflex of the model based on AUC-optimized parameters; on the bottom row – better delineation of the vessel edges and less merging of close vessels by the *Azzopardi method*. The properties are discussed in Subsection 4.3.1.

by the orange lines – the individual profiles – crossing the blue vessel's centreline). For computing the features, six key points were established on each profile – two points (e_l, e_r) corresponding to the vessel edges, two points (h_l, h_r) placed in the upper part of the vessel $0.75 \cdot v_d$ (v_d is vessel diameter) from the middle and two points were placed $0.25 \cdot v_d$ from the middle (c_l, c_r). The points come in pairs, one from the left side and the other from the right side. Furthermore we defined the area between h_l and h_r as the *outer area*, the area between e_l and e_r as the vessel diameter or *central area* and between c_l and c_r as the *inner area*. An illustration of how the points are defined is presented in Figure 4.2. The profile edges were estimated in the green colour channel. The colour channels used to derive the features were taken from the RGB and HSV colour spaces. To eliminate local changes in the image illumination, all channels were normalized by subtracting a median filtered version from itself – the median filter was square shaped with 125-pixel-long sides. The following features were then proposed:

1. Vessel height: the difference in the image intensity at the point $h_{\{l,r\}}$ and minimal intensity in the corresponding profile half; the mean of the feature, computed from the left and right half, is used
2. Central vessel height: the difference between the minimum intensity of the corresponding left or right vessel part is subtracted from the maximum intensity in the *inner area*
3. The ratio between the vessel height and the central vessel height
4. The mean image intensity in the *vessel diameter* area
5. The mean image intensity in the *outer area*
6. The SD of the image intensities in the *vessel diameter* area
7. The SD of the image intensities in the *outer area*
8. The mean intensity under the centreline pixels of the sub-segment in the image smoothed by the median filter
9. Vessel diameter
10. Vessel length

Features 1–7 were established from the red (R), green (G), blue (B), hue (H), saturation (S) and value (V) colour channels and both from the raw channel and the one normalized by subtraction of the median filtered image. Feature 8 was established from each of R, G, B, H, S, V channels. In total, 92 features were proposed.

Greedy forward-feature selection was used to gather the feature set for the final classification. The parameters of the SVM classifier – γ and σ^2 were optimized using the simplex search method of Lagarias et al. [143] (the *fminsearch* function in Matlab was used). Training of the classifier was only done on segments – more precisely the sub-segments used to compute the features – inside the ROI used for estimation of the AVR. The segments selected for classification were limited by diameter >10 pixels, which resulted from preliminary experiments where it was observed that vessels with a smaller



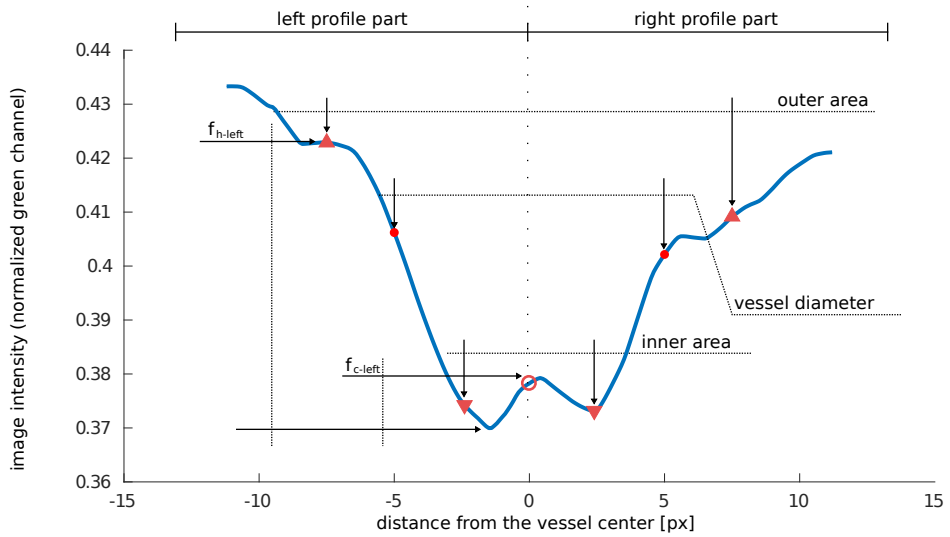


Figure 4.2: An illustration of the profile extraction and estimation of the profile-based features. The vessel segment which is being processed is delineated by the blue solid curve in the upper picture. A sub-segment of the length of 25 pixels is delineated by the brighter orange cross-section lines (every 3rd cross-section is plotted), which corresponds to profiles belonging to the individual sample points of the vessel centreline – these profiles can be seen in the 3D view, colored red. The individual cross-section profiles are then averaged to obtain a single profile from the sub-segment. Filled circles on the profile, detected by the Bankhead algorithm [40], mark the vessel edges delineating the vessel diameter v_d. The circle without any fill in the middle is the mean point between the edge points used to divide the profile's left and right sides. Triangles delineate arbitrarily chosen points – the two labelled *outer area* are estimated as points laying $0.75 \cdot v_d$ from the middle, the points labelled *inner area* are estimated as points laying $0.25 \cdot v_d$ from the middle. The features 1 and 2 from Subsection 4.3.4 are marked as f_{h-left} and f_{c-left}.

diameter are significantly harder to classify. To assign each vessel segment a single label, soft classification of the sub-segments was averaged and used to assign a final label.

For the training process, 40 images were randomly selected from the database and manually labelled as described in Subsection 4.3.3. The set was divided into 20 training images, 10 images used for the feature selection and optimization of the γ and σ^2 parameters (the optimization set) and 10 images were used for the validation of the performance (the validation set) and the estimation of vessel confidence (see Subsection 4.3.4 for details).

Classification confidence

In order to be able to identify vessels that are likely to be classified wrongly, a vessel confidence (v_c), measure was defined. It provides an estimate of how likely a vessel is to be misclassified based on the SVM soft classification. First, the sub-segments located within the ROI were extracted from the images of the validation set and the trained classifier was applied to obtain the set of soft classification values: $S_c = \{s_c^i, i = 1..s_i\}$ and s_i is the number of sub-segments. Then the range of soft classification values was divided into N bins; we used 10 bins in our experiments, and two histograms were obtained from S_c: one for arteries, $H_a = \{h_a^n\}$, and the other for veins, $H_v = \{h_v^n\}$; where $h_\bullet^n$ means the frequency of s_c^i values within the n-th bin. The vessel confidence of an arbitrary vessel segment ρ with the soft classification value s_c^ρ that is closest to the centre of bin k is then estimated as $v_c = 1 - min(h_v^k, h_a^k)/max(h_v^k, h_a^k)$. The confidence was estimated separately for segments thinner than 14 pixels and wider than 14 pixels. A threshold $v_c = 0.85$ was set and vessels with lower confidence (non-confident vessels) were excluded from processing. The threshold was set empirically as a reasonable value to avoid problematic misclassifications.

4.3.5 Automatic estimation of the arterio-venous ratio

For automatic estimation of the AVR, it is necessary to be able to properly select final arteries and veins for computing the ratio. We approached the vessel selection by grouping the vessels based on the smallest angle α_c between their centrelines – groups are created such that each vessel in the group is closer than an arbitrary threshold to at least one other vessel – resulting in the n sets (i.e. groups) of vessel segments, $G_\alpha^i,\ i \in [1..n]$. A second grouping is then done within each G_α^i and is based on vessel parallelity r_p. Paralellity between two vessels is defined by a circular ring with the largest area covering both vessels while none of the vessels has an end point included in the ring; r_p is defined as the thickness of the ring. The second grouping divides G_α^i into m_i sets $G_{pi}^j,\ j \in [1..m_i]$, in this case, all vessel pairs within a set G_{pi}^j have to satisfy $r_p > t_r$ for an arbitrary threshold, t_r; a segments can occur in multiple G_{pi}^j but the intersection of two sets $G_{pi}^k \cap G_{pi}^l$ cannot be equal to either G_{pi}^k or G_{pi}^l. Each vessel left ungrouped forms a new G_{pi}^j.

During the final selection, one set of G_{pi}^j is selected from each G_α^i based on the average distance of the vessels within G_{pi}^j from the OD, where the distance is computed as the shortest distance between the vessel's centreline and the OD's centre. Two approaches to the selection of the final vessels were tested: the first one, AVR_a does not change the

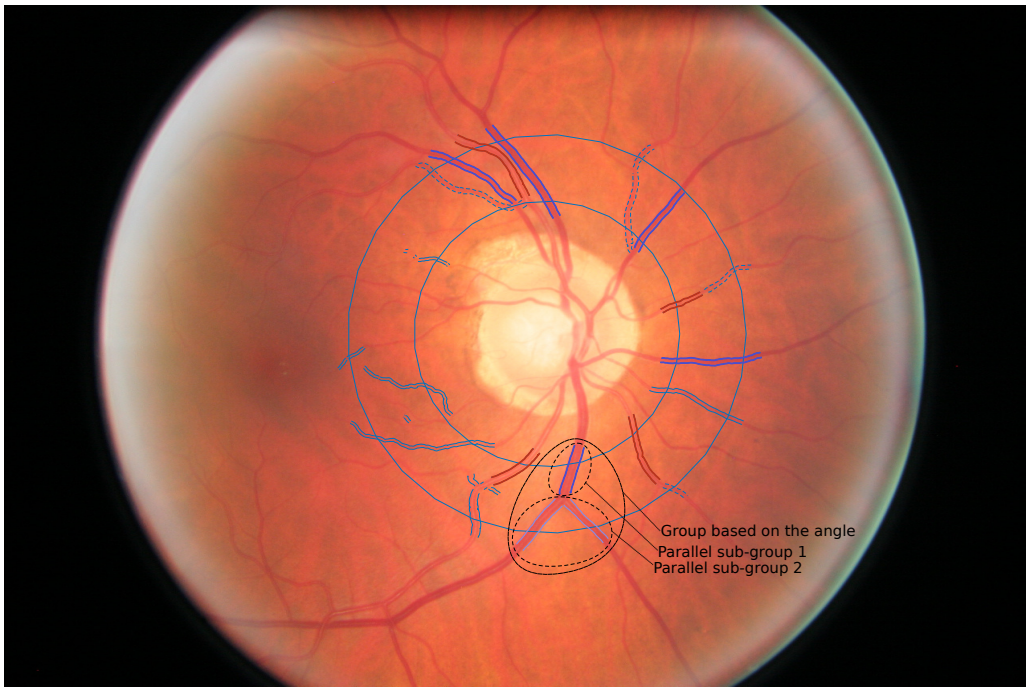


Figure 4.3: An illustration of the final vessel selection. The vessels with delineated edges are the ones which were considered for processing ($v_d > 10$ pixels). The vessels with the bold red edges (classified as arteries) and blue edges (classified as veins) are the final vessels used for AVR estimation. The dashed vessels had classification confidence lower than the selected threshold. The class of vessels that are not the final vessels is not depicted. The three vessels marked *group based on the angle* illustrate the grouping process. The three vessels were grouped by the smallest angle (G_α) and the two subgroups were grouped based on parallelity (G_{pi}). Subgroup 2 (the darker vessel) was selected as the final vessel based on its smaller distance to OD.

final number of either arteries or veins, i.e. when e.g. four final arteries and six final veins are estimated, the numbers are kept as those numbers; the second approach, AVR_l, is such that the number of final vessels is limited to $min(n_a, n_v, 6)$, where n_a is the final number of arteries and n_v is final the number of veins.

When the final number of arteries or veins is lower than four – which can happen, for example, when some of the vessels in the image are estimated as *not-confident* or when many misclassifications happen – the image is excluded from processing.

4.4 Results

The proposed AVR framework is assessed in two ways. First, in Subsection 4.4.1, the accuracy of the vessel segments classification into arteries and veins is presented. The classification accuracy is compared for vessels of different widths and also to the state-of-

Table 4.2: Information on the training, testing and validation sets of the vessel segments. The vessel segments were limited to those inside the AVR ROI with v_d > 10 px and length > 20 px.

	Images	Segments (A, V)	Sub-segm. 25 px
training set	20	410 (220, 190)	2213
test set	10	221 (119, 102)	1197
validation set	10	197 (109, 88)	1070

the-art methods. Then, in Subsection 4.4.2, the associations of the AVR with patients' blood pressure are presented. The association between the AVR and blood pressure is well researched [27] and the associations between the AVR and the blood pressure are used to assess the AVR estimation systems [144].

4.4.1 Accuracy of the arterio-venous classification

Assessment of the AV classification was done by assessing the accuracy, sensitivity and specificity of the centreline pixel classification. Accuracy is the percentage of correctly classified pixels from all the tested centreline pixels. Sensitivity is the percentage of correctly classified pixels labelled as arteries in the GT. Specificity is the percentage of correctly classified pixels labelled as veins in the GT (specificity). At the same time, performance of the classification among the whole segments was assessed because each segment comes to the AVR computation process as one piece, providing a single diameter value. Also, the percentage of correctly classified segments was used as a measure for feature selection.

The segments included in the process were filtered by their diameter and length, and only segments inside or crossing the ROI were taken into account. The width and length limits were such that only segments wider than 10 pixels and longer than 20 pixels were considered. The length p_n of the sub-segments was set to 25 pixels. Final classification of a vessel was then obtained as an average soft label of the sub-segments within the vessel. Statistics on the properties of the vessel segments is provided in Table 4.2.

Feature selection done on the test set resulted in the feature set $\{Rn_3, Gn_5, B_2, G_2\}$ (a capital letter marks the colour channel, n after the channel indication means normalized image and the lower index marks the feature number as defined in Subsection 4.3.4). An assessment of the AV classification performance is in Table 4.3. Pixel-wise accuracy of the AV segments that were wider than 10 pixels was 0.96 on the validation set with the performance skewed significantly between narrower and wider segments, as segments with 10–14 pixels reached an accuracy of 0.93 and segments wider than 14 pixels reached 0.998. However, the drawbacks are that when non-vessel segments appear in the image, they are classified as either an artery or vein because the proposed system does not consider any other classes. When the non-vessel segments are included in the assessment, the accuracy declined to 0.946, 0.928 and 0.998 respectively. In reality, the non-vessel segments are typically shallow and can be confidently classified as arteries.

Segment-based accuracy was lower than that of the pixel-based one, sliding to 0.89 in the case of non-vessel inclusion. For the final AVR estimation system, it is important

Table 4.3: The accuracy of the final classifications (both pixel-based and segment-based classifications). Division on the narrow and wide segments (bottom part) show much better performance of the classification for the wider vessels, that are more important in AVR estimation. An assessment was also made solely on AV segments and on AV segments with non-vessels (n-v) included.

			Accuracy		Sensitivity		Specificity	
Diameter	N		Centerline	Segments	Centerline	Segments	Centerline	Segments
>10	197	AV only	0.964	0.934	0.998	0.991	0.930	0.864
	208	incl. n-v	0.946	0.885	0.960	0.90	0.894	0.768
10–14	100	AV only	0.928	1.0	0.822	0.880	1.0	0.667
	111	incl. n-v	0.891	0.793	0.935	0.853	0.746	0.511
>14	97	AV only	0.998	0.990	0.996	0.978	1.0	1.0
	97	incl. n-v	0.998	0.990	0.996	0.978	1.0	1.0

that the specificity (the true positive rate for veins) in the segment-based case is 0.77, which means 23 % of the wrongly classified vein segments.

In comparison with the state of the art, the classification performance is among the best state-of-the-art methods. Typical reported accuracy of the AV classification is around 90 %. Grisan et al. [49] reported accuracy of the segment-based classification of 88.6 % of 443 segments obtained from 24 images with an accuracy of 93.7 % of the main vessels (vessels belonging to the main vessel arcs). Konderman et al. [53] reported pixels-based accuracy of 95.3 % with no indication on the vessel width limits. Saez et al. [55] reported the accuracy of artery classification (equivalent to our sensitivity) to be 79.2 % and the accuracy for veins (equivalent to our specificity) to be 87.9 %. Niemeijer et al. [57] reported pixel-based accuracy of 84 %.

The approach most similar to this work was proposed by Relan el al. [44]. The study concluded pixel-based accuracy of 94.88 %, equivalent to our sensitivity of 97.9 % and equivalent to our specificity 90.1 %. The exact width of the vessels included in the computations cannot be compared between the approaches but considering their limitation of using vessels wider than 55 % of the maximal width in pairs of quadrants – the mean maximal width per image in our approach was 24.7 pixels, leading to 13.6 pixels as the minimal width limit (this estimation is the upper bound of Relan's limit because it does not consider the quadrant-wise computation) which is higher then out limit 10 pixels. Considering that Relan et al. used 802 vessel from 70 images, the numbers are consistent with our 394 vessel segments from 40 images. The results obtained in our study are thus considered comparable to (or considered to outperform) other state-of-the-art feature-based approaches.

4.4.2 Arterio-venous ratio and its associations

From the 167 testing images, 35 were excluded for being invalid during the automatic processing. The reason for the exclusion was typically low classification confidence of

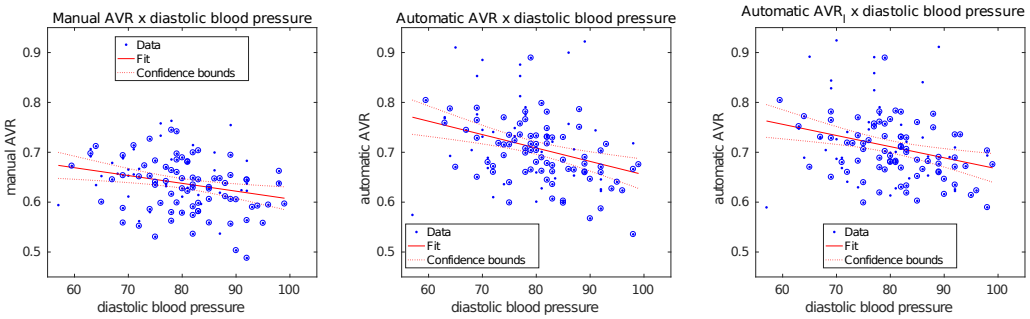


Figure 4.4: Linear models between the AVR and diastolic blood pressure. The single dots are all the results of all automatically segmented images (without those automatically excluded) and the dots with circles are the data points gathered from the correctly classified images. The left axes show the manually estimated AVR, the middle axes show AVR_a and the right axes show AVR_l.

several vessels while a few misclassifications occurred at the same time, which results in one vessel class having less than four final vessels. The results from the other 132 images were used to estimate linear model M_1 between the AVR and blood pressure (Figures 4.4 and 4.5). These images were manually divided into these still containing misclassified vessels and those where all vessels were classified correctly. Another 43 images were sorted out during the process which left 89 test images. These were used to create linear models M_2 and to estimate the influence of the wrongly classified vessels on the resulting AVR.

Associations between the manual and automatic AVR and diastolic and systolic blood pressure can be seen in Figure 4.4 and 4.5. A summary of M_1 and M_2 is provided in Table 4.4. As shown in the figures and also depicted in the table, both automatic and manual AVRs were negatively correlated with diastolic blood pressure; the slope of the fitted linear models and their statistical significance were stronger in the case of the automatic procedures and strongest in the case of AVR_a. Model M_2 showed a slight increase in the significance of the parameters versus model M_1. The presence of the misclassified vessels had an influence on the root mean square error (RMSE) of the models. This can be seen clearly in the figures where outliers are present, especially above the linear model, although it should be noted that only about seven of these outliers lie significantly out of the range of correctly classified vessels. The RMSE values of model M_2 are very close between the models AVR_m, AVR_a and AVR_l.

Associations with the systolic pressure were observed only for the automatic AVR and only for model M_2. The variation in the results was also close to that in the case of diastolic pressure. From the other clinical parameters, BMI was the only parameter found to be significantly associated with the AVR. Age and sex were not found to be significantly associated with the study sample.

Correlation coefficient between the manual and automatic AVR, when all 132 images were considered, was 0.48 for AVR_a and 0.52 for AVR_l. When those 89 correctly classified images were considered, the correlation coefficient was 0.53 for AVR_a and 0.59 for AVR_l. All the correlations were significant ($p < 10^{-6}$).

Table 4.4: Associations of the estimated AVR with systolic and diastolic blood pressure. AVR_m marks AVR as estimated by manual grading, AVR_a is AVR as estimated by the automatic approach and the final vessels are included in the computation without restrictions and AVR_l is AVR where the final vessels were limited so that amount of veins and arteries was equal. M_1 is the model where all automatically assessed images were included and M_2 is the model where only images without misclassifications were included. I is the model's intercept and L is the model's linear coefficient. Linear coefficients were multiplied by 10 for convenience, i.e. L = -0.027 means average change of the AVR with increase of 10 mmHg in the blood pressure. CI is confidence interval.

Diastolic pressure

		AVR_m			AVR_a			AVR_l		
		coef.	CI	RMSE	coef.	CI	RMSE	coef.	CI	RMSE
M_1	I	0.76	[0.68, 0.84]	0.053	0.92	[0.81, 1.04]	0.07	0.89	[0.78, 1.00]	0.067
	L	-0.015	[-0.026, -0.005]		-0.027	[-0.041, -0.013]		-0.022	[-0.036, -0.009]	
M_2	I	0.78	[0.67, 0.90]	0.052	0.96	[0.85, 1.08]	0.054	0.9	[0.79, 1.01]	0.052
	L	-0.019	[-0.033, -0.005]		-0.033	[-0.047, -0.019]		-0.025	[-0.039, -0.011]	

Systolic pressure

		AVR_m			AVR_a			AVR_l		
		coef.	CI	RMSE	coef.	CI	RMSE	coef.	CI	RMSE
M_1	I	0.71	[0.62, 0.8]	0.055	0.8	[0.68, 0.92]	0.074	0.78	[0.67, 0.89]	0.07
	L	-0.005	[-0.012, 0.001]		-0.007	[-0.016, 0.002]		-0.005	[-0.014, 0.003]	
M_2	I	0.73	[0.62, 0.84]	0.054	0.85	[0.73, 0.97]	0.058	0.82	[0.79, 1.01]	0.055
	L	-0.007	[-0.016, 0.001]		-0.012	[-0.021, -0.003]		-0.009	$[-0.017, -1 \cdot 10^{-5}]$	

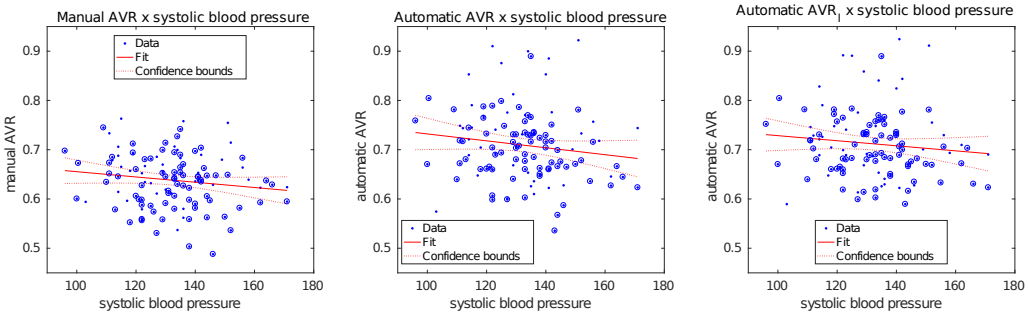


Figure 4.5: Linear models between the AVR and systolic blood pressure. The single dots are all the results on all automatically segmented images (not the automatically excluded ones) and the dots with circles are the data points gathered from the correctly classified images. The left axes show the manually estimated AVR, the middle axes show AVR_a and the right axes show AVR_l.

4.5 Discussion

The system we use to predict the parameters of the vessel segmentation methods is proposed for the first time to best of our knowledge. Although only qualitatively assessed, it was concluded that the segmented vessels were obtained with sufficient quality. Although comparison with manual vessel segmentation would be of great interest, the topic itself is not straightforward because only a small amount of research [145, 146, 142] has focused on the selection of proper methods for the assessment of binary vessel segmentation.

Proper assessment of the vessel segmentation is a complex topic owing to the fact that it is dependent on its application and that the vessel edges are not defined in the objective way in the typical GT. For example, most research assesses the proposed methods for vessel segmentation using pixel-wise accuracy – as was done in Chapter 3 of this work – compared with manual segmentation. Considering only the centerlines of the vessels is important in the subsequent vessel tracing using the *Bankhead method*; it has to be concluded that the optimized parameters of vessel segmentation might only be rough estimates of the optimal values in the application sense. It was concluded, based on our results, that when the segmentation parameters were optimized by AUC, the resulting vessel segmentation was more robust, and in applications similar to the one presented here it is encouraged that researchers consider AUC as the measure for comparison of vessel segmentation performance.

The proposed features for AV classification were mostly based on the state of the art and did not broaden the knowledge about the applicable features. The set of features selected by feature selection confirms the strong predictability of the contrast-related features: the ratio of the height of the central reflex and the height of the whole vessel. The height of the central vessel reflex, which was among the selected features, is a novel method considering the literature reviewed for the purpose of this work. Also it was concluded that the profile-based features are as strong as or better than features based on the neighbourhood of the centreline pixels, which were employed in most of the recent studies. More direct comparison will be needed in order to be able to draw conclusive results indicating the quality of the features.

The confidence of the vessel classification, as defined in the presented work, has not been presented elsewhere. The concept of an *undecided* classification is not new and has been applied in other approaches, by Relan et al. [45] for example. Approaches that rely on unsupervised classification typically define an interval or area in the feature space that is considered uncertain. The approach proposed in the presented work is novel in the usage of the probability estimate for misclassifications. The probability estimate could have an important impact in the future application of structure-based AV classification, where the confidently classified labels can be propagated through the vessel network to the segments with lower confidence. In our case, it allowed us to estimate the quality of the classification of images where the AV GT is not present and to optimize the necessary number of training images or other parameters.

The proposed approach to the selection of the final vessels offers an alternative to the method which averages the AVR estimated from several sub-rings with small thickness. The proposed approach was not compared to this method so it is not possible to conclude how the results differ. The advantage of the proposed approach is that it provides the

possibility to decide to include either the branches or roots of the vessels within the AVR computation. Compared to other approaches, like tracing the vessel graph, this approach can be less prone to errors in the graph estimation, such as vessel disconnections. The grouping of the vessels is not computationally expensive unless some groups contain many vessels before the parallelity grouping is performed. In that case, the computation can be problematic because all the combinations of the vessels are tested. This situation does not happen often though, and in our experiments the image was excluded from processing when such a case emerged.

Statistical analysis of the automatically estimated AVR gave very promising results when a stronger association between the blood pressure and the automatically estimated AVR was concluded. Furthermore, a stronger association was found between the AVR computed without regularization of the number of final vessels in the sense of limiting the number of either arteries or veins when a number of vessels in one class exceeded the other. This result could be explained by the fact that when a vessel of one type branches less often than the other, all vessels from the more-branched vessel should be included in the computation. When the number is limited it actually raises a bias. Although there was a relatively high number of images with misclassifications (33 %), approximately only around six vessel images with misclassified vessels resulted in outliers in the final AVR estimation. The influence of the misclassifications in the other images was most likely not that significant because only narrow vessels were misclassified.

It was concluded that the automatic approach can provide an advantage over manual measurement of the AVR by showing stronger associations with the clinical measurements. The fact that the manual measurements were established independently and before the presented study started supports the validity of the results. The manual measurements were done in the pure manual way, without employment of a semi-automatic approach (like, for example, that used in [144]). There could be important differences when such approaches are employed and further research should focus on addressing such questions.

The proposed automatic AVR estimation framework has the advantages of the automatic prediction of the vessel segmentation parameters, a simple and robust method for final vessel selection and an AV classification approach comparable with the state-of-the-art approaches. The disadvantages are the missing automatic detection of the OD, the supervised nature of the AV classification and no employment of the vessel structure in support of the vessel classification. Those are the main weaknesses that should be addressed in the future.

Chapter V

Processing of the diaphragm image sequences

5.1 Introduction

In this chapter, methods for the assessment of a non-respiratory diaphragm function are presented. Screening of the diaphragm was done by functional MRI. The main goal was to separate respiratory diaphragm movements from non-respiratory movements, and then to evaluate their role in body stabilization. The subjects included in the study consisted of a group of healthy volunteers, and a group of volunteers in whom structural spine disorders had been identified and who suffered from LBP that had lasted for at least one year.

The diaphragm was investigated to assess reactability and movement during tidal breathing and breathing while a load was applied to the lower limbs. The harmonicity of the diaphragm movement, frequency and range were used to characterize both the respiratory and postural movements. Another part of the assessment was acquired from the static diaphragm, where the inclination, height and position in the trunk were estimated. Two methods were applied for the estimation of diaphragm movement. One was based on a similar method [109] and the other was based on the registration of the blood vessels visible caudally from the diaphragm in the MRI pictures. The differences between healthy subjects and subjects with LBP were evaluated statistically. The results of our work should help in understanding the function of the diaphragm in the posture stabilization system, with possible implications for physiological practice.

5.2 Study settings and data

5.2.1 Study settings

Diaphragm activity was monitored under two different situations:

- S_1 – subjects lie supine on their backs during tidal breathing.

Table 5.1: Details of the study groups (mean ± SD). C_1 is the control group; C_2 is the patients group.

	C_1	C_2
Age	35 ± 11	42 ± 11
Weight (kg)	71 ± 15	78 ± 16
Height (cm)	172 ± 10	174 ± 6
Sternum height (cm)	20.9 ± 1.61	21.4 ± 1.77
Thorax height (cm)	30 ± 2.1	30.2 ± 1.7

- S_2 – subjects lie supine on their backs during tidal breathing while loading is applied to the distal part of their extended lower extremities against the flexion of the hips. The applied pressure was of the fourth grade according to Janda's muscle test [147]. The subjects ensured that no additional pain was induced by the maneuver.

5.2.2 Subjects groups

Two groups of volunteers were selected:

- C_1 were subjects without a pathological condition (n=16; 11 women, 5 men: the control group) and were given id numbers 1–16.
- C_2 were subjects with a structural pathological condition of the spine localized in the lumbar spine area (n=17; 8 women, 9 men: the patients group) and were given the id numbers 17–33.

Neither the control group nor the patients group had any pulmonary disease. The average age of the control group was 35 years old (in a 23–56 age span). The average age in the patients group was 42 years old (in a 23–65 age span). Detailed characteristics of the two groups are summarized in Table 5.1.

Structural findings in the patients group were confirmed by the previous MRI in the lumbar spine area. The study excluded patients with an inborn defect of the spine or a defect acquired traumatically. All patients group had suffered from LBP of various intensity and frequency for at least one year (classifying the LBP as chronic). The intensity of the LBP was determined using a visual analog scale (VAS) with a range of 0–10. The subjects indicated their current pain on the day of imaging and their overall pain over the course of one month before imaging. The VAS values for the control group C_1 were zeros for all subjects.

Acute pain was not the criterion for selection of the patients group. The main criterion was the spinal findings.

Due to a distinct inter-group difference in age, a paired t-test was performed to confirm no statistical significance among the groups. The resulting p-value ($p = 0.08$) showed no statistical difference at the 5 % significance level. The normality of the age distribution within the groups using the Kolmogorov-Smirnov test ($p_{C1} = 0.89$, $p_{C2} = 0.55$) confirmed

normal distribution of the data. No differences in the mean for all other parameters in Table 5.1 were confirmed with great significance ($p > 0.2$).

The other portion of the data studied in this work consisted of healthy, asthma and COPD groups of patients. These data were processed in the study [20]. The dataset consisted of images from 31 patients measured in situations S_1 and S_2.The apparatus for the data acquisition was the same as in the case of group C_2 (see Subsection 5.2.3). This data portion was used to improve the method for motion estimation of the diaphragm as reported in Subsection 5.3.5.

Ethics statement

All patients provided written, informed consent for participation in the study. The study was approved by the Ethics Committee of the Motol University Hospital in Prague, Czech Republic.

5.2.3 Data acquisition

The healthy group was examined in an open Siemens MRI apparatus, with a 0.23 T magnet and the NUMARIS/4 syngo MRI 2004A software version. The length of each recorded sequence was 18 s. During this time, 77 images were recorded at regular intervals. The subject was in a supine position, using a large body-size coil. The projection plane was placed sagittally in the axial topogram and directed paravertebrally down the right side, midway through the center of the vertebral body and the edge of the thoracic wall. The width of each layer was 33 mm. The true FISP dynamic sequence was used, configured as follows: 1 NSA, matrix 240x256 pixels, TR 4.48 ms, TE 2.24 ms, FA 90, TSE1, FOV 328 mm.

The patients group was scanned by General Electric SIGNA HDx MRI, with a 1.5 T magnet and the 14-M5A software version. The length of each recorded sequence was 22.2 s, resulting in the acquisition of 60 images. The projection plane was placed sagittally in the axial topogram and directed paravertebrally down the right side, midway through the center of the vertebral body and the edge of the thoracic wall. The width of each layer was 15 mm. A GE FIESTA cine dynamic sequence was used, configured as follows: 1 NSA, matrix 256x256 pixels, TR 3.1 ms, TE 1.3 ms, FA 55, FOV 420 mm.

The proposed processing methodology is independent on distinct images resolution, e.g. the resolution of the control group was 1.37 mm/pixel and the resolution of the patients group was 1.64 mm/pixel. Three markers, 20-ml syringes filled with water, were placed on the skin surface of each subject on her or his right side. They are shown as hyper signal marks on the body surface (see for example Figure 5.1). The first marker was placed in the mid-clavicular line at the level of the jugular notch, and the second marker was placed at the level of the inferior 10-rib costal margin. The last marker was placed on the back of the subject at the level of the thoracolumbar junction.

The spatial resolution of our images was sufficient for proper diaphragm contour recognition on each sequence. In addition, the temporal resolution was sufficient, with a maximum recorded breathing frequency of 0.54 Hz. The diaphragm contour areas were not affected by artifacts. Image brightness was the only varying parameter, and there was no significant effect on the resulting differential curves.

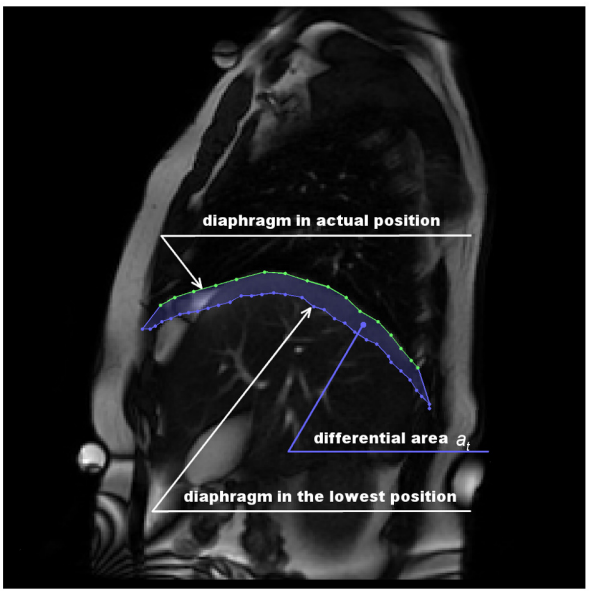


Figure 5.1: Differential area definition. Figure shows t-th image from a sequence with the correnponding diaphragm contour. The t-th diaphragm contour together with the lowest placed diaphragm contour in the sequence delineate the differential area a_t.

5.3 Methods

5.3.1 Processing of the input images

In order to assess diaphragm activity, the differential curve (diff-curve) was calculated across all MR images. Firstly, let us define the differential area (a_t) as the area bordered by the diaphragm in the lowest position from the sequence and the diaphragm in current (t-th) image – see Figure 5.1. The image containing the lowest placed diaphragm is called the background picture. Secondly, the diff-curve is defined as the time series of a_t, measured in mm^2. Hence, the diff-curve is an integral quantity which characterizes the diaphragm motion in the same manner as spirometry, but it consists strictly of diaphragm movement. The algorithm for a_t calculation is shown in Figure 5.2A–C. Typical diff-curves are shown in Figures 5.3 and 5.4.

5.3.2 The extraction of respiratory and postural movements from diff-curves.

Each digitally sampled signal can be expressed as the sum of a finite number of harmonic waves of different amplitudes and phases. Discrete decomposition of the signal into harmonic components is traditionally represented by the harmonic spectrum of the signal. The spectrum denotes the relation between the amplitudes and frequencies of the harmonic waves. Typical diff-curve spectra are shown in Figure 5.3. Each peak in the

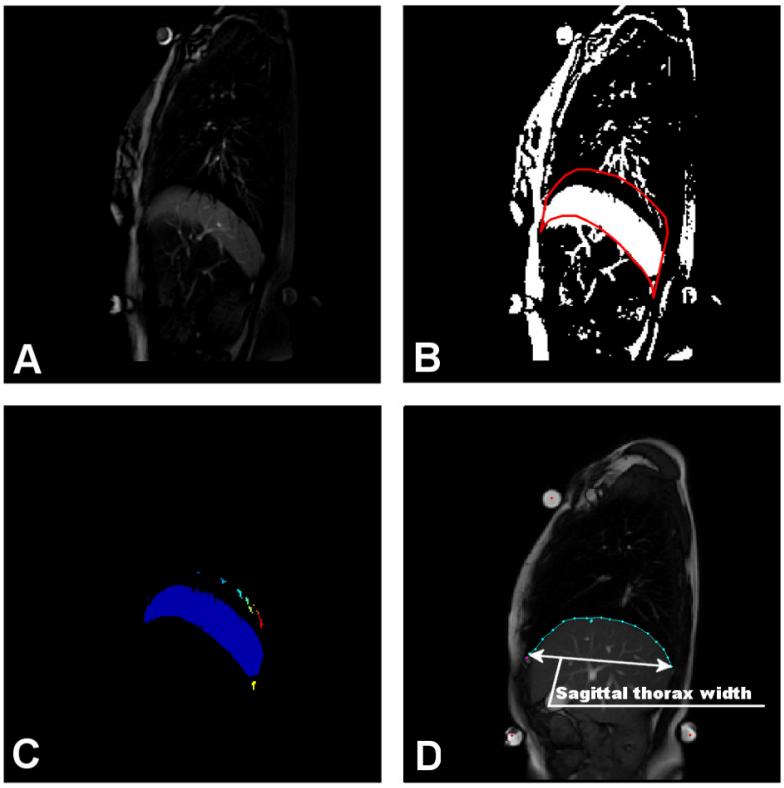


Figure 5.2: Differential area calculation. The background image (the image with the lowest placed diaphragm) is subtracted from the image on t-th position in a sequence (A). The subtracted image is thresholded, providing a binary image with a clearly visible crescent corresponding to the movement of the diaphragm (B). The red-bordered part surrounding the highest and the lowest diaphragm position from the whole sequence. Continuous image parts inside the border are labeled and the part corresponding to the diaphragm movement is processed (C). Some of the extracted parameters were normalized using the thorax width (D).

spectrum represents a single harmonic component. If the diaphragm only worked for respiration with stable depth of the motion, it would lead to a simple spectrum with a single peak corresponding to the breathing frequency. This motion would be fully described by a single sine wave. Diaphragm motion is more complex, and often involves other, non-respiratory movements. However, due to the harmonic properties of respiration, the harmonic spectrum is useful for diff-curve processing.

We chose to model diaphragm motion using two sine waves corresponding to the respiratory and non-respiratory movements. These waves are extracted from the diff-curve spectrum using the inverse Fourier transform. The two models are shown in Figure 5.3 A and B. The original diff-curve is plotted by a solid line. The respiratory model is called a respiratory curve (*res-curve*). The non-respiratory model is called a postural curve (*pos-curve*). The *res-curve* fully characterizes the respiratory movement by frequency and amplitude. The *pos-curve* provides a model of a postural global range using a *pos-curve* amplitude. If there were several peaks that together compose the final respiratory or postural part of the signal, we chose the peak that best described the values of the original signal (peaks and subsequent *pos-curves* are marked in Figure 5.3 with a green square).

The relation between breath regularity and the spectrum complexity of the corresponding diff-curves is summarized in Figure 5.3. Figure 5.3 A and C, provides an example of a diff-curve with the corresponding spectrum for a person whose respiratory movements are very regular. In the corresponding spectrum, there is one significant peak, which represents the subject's respiration marked by a red dot. The diaphragm's respiratory movement is also modulated by other movements. This causes the occurrence of smaller peaks besides the respiratory peak. These peaks capture smaller parts of the diaphragm's movement. Figure 5.3 B and D show a more complex diff-curve with less regular respiration. The spectrum (Figure 5.3 D) has a clearly visible peak, which corresponds to the respiration (marked by a red dot) and, again at lower frequencies, there are peaks which modulate the respiratory movement. This time, however, the peaks capture a much larger part of the diaphragm's movement.

Three typical diff-curves with corresponding respiratory and postural models for both situations (S_1 and S_2) are shown in Figure 5.4. There is a clearly visible respiratory function (A) with a big postural movement in situation S_2 (B) for subject 11. Subject 25 (in E, F) did not respire for the first six seconds of the imaging in situation S_2, and then his or her respiration became regular. Subject 24 (C, D) exhibited almost no respiration during situation S_2. Diff-curves including no respiration led to exclusion of the subject from further statistical data processing.

5.3.3 Parameter extraction

Two sets of parameters were extracted on the basis of diaphragm MRI activity – dynamic parameters and static parameters. The dynamic parameters are based on the processing of the diff-curve. The main aim of introducing the dynamic parameters was to assess which part of diaphragm motion is related to respiration, and how significant the non-respiratory movements are. The dynamic parameters are as follows:

- Frequency (f_r) and amplitude (a_r) of the *res-curve*.

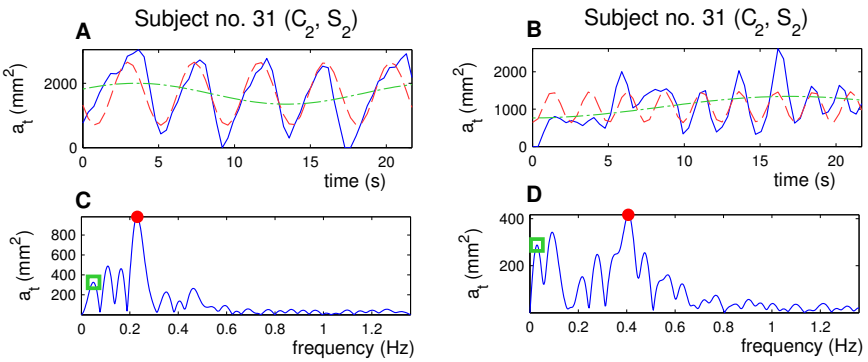


Figure 5.3: Diff-curves (A, B, the solid line) and appropriate spectra (C, D, the solid line). Extracted *res-curves* (the red dashed line, A, B) and *pos-curves* (the green dotted line, A, B) with corresponding spectral peaks (C, D) marked in the spectra with a red dot (the respiratory peak) and a green square (the postural peak).

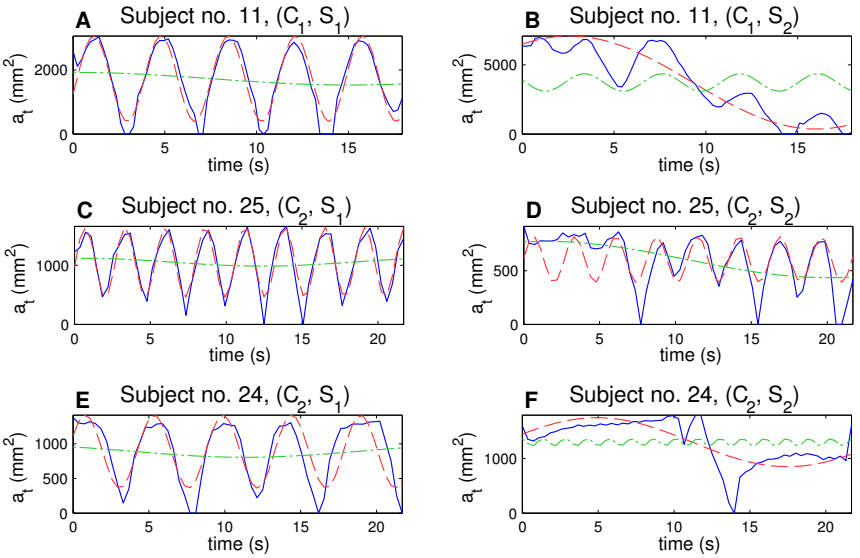


Figure 5.4: Diff-curves (the solid line) and extracted *res-curves* (the red dashed line) and *pos-curves* (the green dotted line). An example of harmonic breathing (A), breath with a strong postural part after the load occurred (B), harmonic breath which became partly non-harmonic after the load occurred (C, D), and breath which almost lost its ability for respiratory movement after the load occurred (E, F).

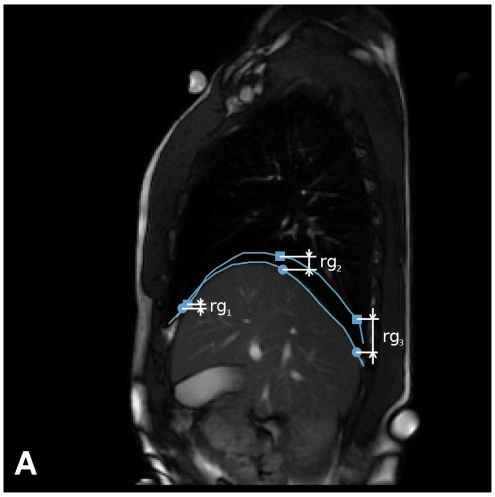


Figure 5.5: Parameters rg_i: Parameters rg_i are computed as a vertical subtraction of the caudal diaphragm position from cranial diaphragm position. The three parameters correspond to the anterior (rg_1), middle (rg_2) and posterior (rg_3) diaphragm parts. Points were spread evenly on the diaphragm contour with a small constant drift of rg_1 and rg_3 from the contour margins.

- Amplitude of the *pos-curve* (a_p).
- The amplitude ratio of the *res-curve* and *pos-curve* ($r_{pr} = \frac{a_p}{a_r}$).
- Range of the diaphragm motion measured in mm in three different points (Figure 5.5) placed on diaphragm surface ($rg_i, i \in 1, 2, 3$).
- The percentage of energy represented by the three highest spectrum lines (p_3).
- The SD (σ_{DC}), skewness (γ_{DC}) and kurtosis (β_{DC}) of the diff-curve.

The static parameters assess the shape and position of the diaphragm. For the static parameters, the following features were extracted:

- The inclination of the diaphragm in the sagittal plane in the caudal position (dec_a); the angle measured as shown in Figure 5.6
- The distance of two lines parallel with the line used for estimation of dec_a and intersecting the most caudal and most cranial points (Figure 5.7) of the diaphragm contour (h_d)
- The vertical distance from the anterior point used for rg_3 and back marker (d_p); this parameter corresponds to the diaphragm height in the thorax (Figure 5.8)

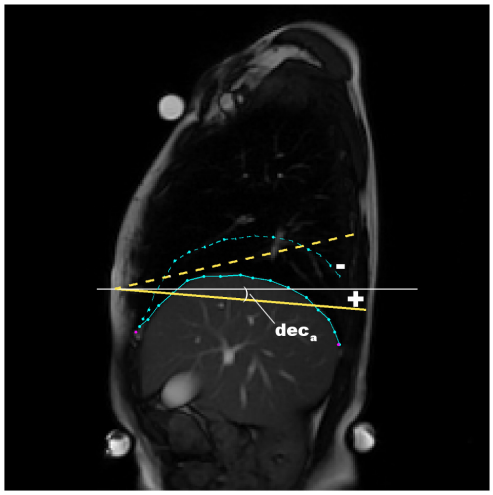


Figure 5.6: Measurement of the diaphragm inclination. The inclination of the diaphragm was measured by the angle between the line fitted to the diaphragm contour and horizontal axis. The inclination was measured during the caudal diaphragm position.

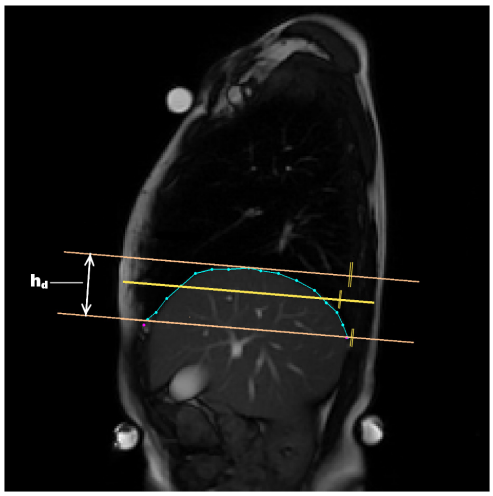


Figure 5.7: Measurement of the diaphragm height h_d. Measurement was done during the diaphragm caudal position. The middle line is the line fitted to the diaphragm contour by the least squares method.

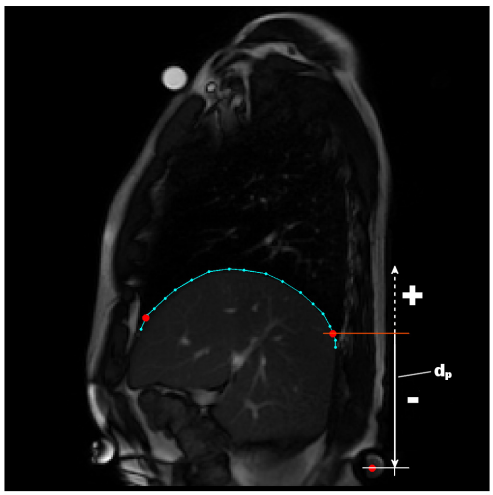


Figure 5.8: Measurement of the diaphragm height in the trunk. The figure indicates when the height is negative (the diaphragm higher than the back marker) and when the height is positive (the diaphragm is placed under the back marker).

5.3.4 Statistical analysis

A paired t-test was used to identify differences between the control group and the patients group. The significance of the statistical test is marked * ($p < 0.05$), or by the symbol ** ($p < 0.001$) as indicated in Tables 5.2–5.4 in Section 5.4. A Kolmogorov-Smirnov analysis was performed to assess the normality of the data.

The correlation (by Pearson's correlation coefficient) between all the parameters and the subjects BMI was assessed in order to eliminate any possible effect on the results. The parameters affected by the BMI dependence were in situation S_1: p_3 and h_d and in situation S_2: $a_r, \beta_{DC}, \gamma_{DC}$ and d_p. A possible way to suppress the correlation with the BMI is to normalize the parameters by the width of the subject's thorax (Figure 5.2 E; the width was determined during the lowest position of the diaphragm). However, no influence on the statistical results was observed after normalization, except for the parameter h_d, for which the results of the Student's t-test changed by two orders of magnitude (but there was no change of significance). In order to keep the results clear, all parameters were kept in the original units, with the exception of h_d, which is in the normalized form.

All the extracted features were treated for outlier values. Outlier values were determined as follows:

$$\text{proper data range} = [P_{25} - w \cdot (P_{75} - P_{25}), P_{75} + w \cdot (P_{75} - P_{25})] \tag{5.1}$$

where P_k means k-th percentile, w is a constant set by default to 1.5 (ensuring approximately 99.3 percent coverage of the data when the data is normally distributed). Data outside this range is likely to consist of error values or marginal data that distorts the statistics.

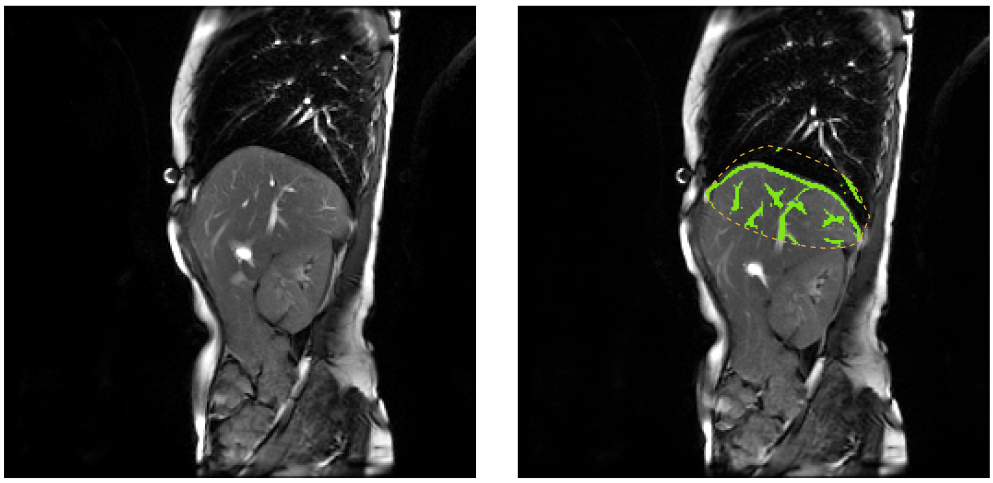


Figure 5.9: An MRI image of the diaphragm (on the left), the corresponding vessels (on the right, green coloured) and the ROI (the dashed curve).

Secondly, as stated in the methodology section (Subsection 5.3.2), patients were present in our datasets whose respiration did not exhibit proper respiration movement. These subjects were also excluded from the statistical evaluation – four subjects, all from the patients group C_2 (id numbers: 19, 24, 27, 29).

5.3.5 Enhanced extraction of the diaphragm movement

In order to provide estimation of the diaphragm movement in arbitrary points on the surface of the diaphragm, a method was established to track the motion of the diaphragm. Vessels that are visible located caudally from the diaphragm contour, were segmented and used as a landmarks for registration. Extraction of the vessels was done by the Nguyen method which has been described in Chapter 3. The method was selected based on preliminary experiments. The methods used for the retinal blood vessel segmentation, presented in Chapter 3, were already well tested and understood by the author so selection was limited on these methods. The *Azzopardi and Nguyen methods* provided better and more versatile results then the *Bankhead method* and were unsupervised which was necessary condition for application in this case. The Nguyen method then provided good segmentation results while setting the method parameters is much more straightforward than in the case of the *Azzopardi method*.

The segmented vessels were limited to a ROI delineated by the diaphragm contour in the most cranial position and an half-ellipse which was attached to the endpoints of the diaphragm contour in the most caudal position. The eccentricity of the ellipse was chosen as 0.3. Both the detected vessels and the ROI are illustrated in Figure 5.9.

After the vessel detection the vessel structures were registered on each other between the successive images in the image sequence. Rigid transformation, employing translation and rotation of the segmented vessels, was chosen for the registration. Selection

of the transformation type comes from the nature of the task where the movement of the whole structure is of interest, small displacement of the individual vessels are actually causing bias. The objective function of the registration process was defined as $\sum_{pixels} |v_{s1} - v_{s2}|$, where v_{si} is either binary or soft vessel classification. The registration parameters – translation and rotation – were found by a combination of grid search over the binary segmentation (to obtain preliminary translation parameters) and the simplex search method of Lagarias et al. [143] where the translation obtained by the grid search was used as the initial value and rotation was added among the minimized parameters. During the grid search procedure binary segmentation was used as the v_{si} and during the fine-tuning process, soft classification was used. The registration process resulted in a transformation T such that $T(v_1) = v2$, where v_i refers to the vessel classification in the i-th image. The pair-wise registration of the images in the sequence then resulted in set of transformations $T_a = \{T_1..T_{n-1}\}$, where T_i is transformation of the i-th image on (i+1)-th. The registration process is illustrated in Figure 5.10.

The set T_a can be used to directly obtain the motion of an arbitrary point, e.g., on the diaphragm's contour. The motion of the diaphragm when estimated this way has the advantage of providing information not only about the cephalo-caudal diaphragm motion, but also about its ventro-dorsal motion. Furthermore, estimation of the cephalo-caudal motion is not biased by the arc-like shape of the diaphragm in the dorsal and ventral parts. The outcome of the motion estimation is illustrated in Figure 5.11 where three points – in the ventral part of the diaphragm, in the middle and in the dorsal part – were selected and their motion compared. In the figure, the motion range of different points differ, which suggests that the diaphragm was rotating during the motion and at the same time the ventro-dorsal movement has significant range when compared to the cranio-caudal direction and assessing only the cranio-caudal range could lead to misinterpretations.

5.4 Results

The results of the dynamic parameter measurements and comparison are presented in Subsection 5.4.1 and in Tables 5.2 and 5.3. The results of the static parameter measurements and comparison are presented in Subsection 5.4.2 and Table 5.4.

5.4.1 Dynamic parameters

Respiratory and postural curves

We observed significantly faster respiration in the patients group in both observed situations (S_1, S_2), with $p < 0.05$. Respiratory frequency did not change much for the control group after a load was applied to the lower limbs (0.21 Hz in S_1, 0.22 Hz in S_2). By contrast, the frequency of the patients group rose significantly ($p = 0.01$). The height of the diaphragm respiratory movements reflected by the respiratory curve amplitude (a_r) resulted in a very significant difference among the groups, with $p < 0.001$ in both situations S_1 and S_2. As in the case of respiratory frequency, there was no change in respiratory curve amplitude in the control group when a the load was applied to the

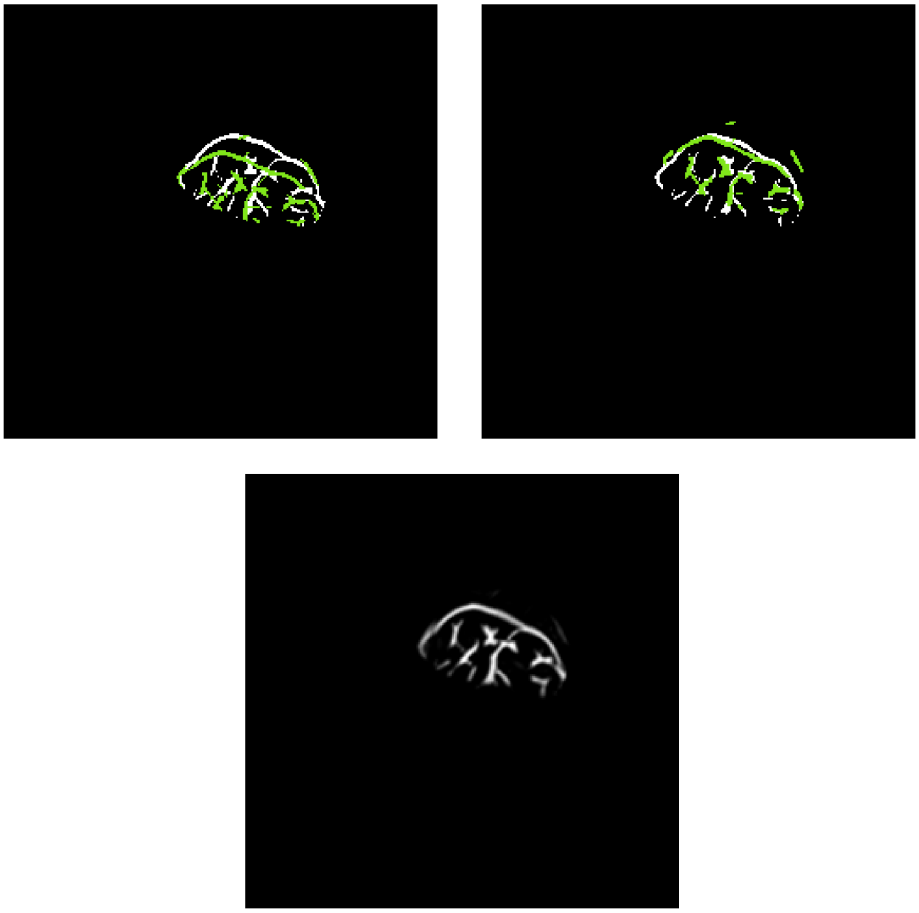


Figure 5.10: The registration of the diaphragm. Two vessel structures are visualized in their original position (the left image) and after the registration (the right image). The picture at the bottom shows the averaged vessel image from all 60 images after registration.

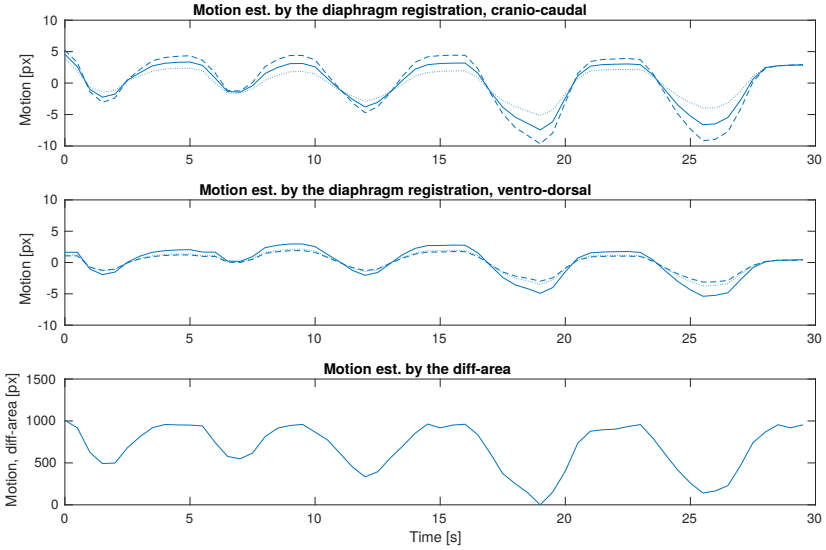


Figure 5.11: A comparison of the diaphragm motion estimated by the registration of the vessel structures (top and middle axes) and by computation of the diff-curve (bottom axes). In the axes corresponding to the registration-based motion, the dashed line corresponds to the most ventral diaphragm part, the solid line corresponds to the middle part and the dotted line corresponds to the dorsal part of the diaphragm.

Table 5.2: Dynamic parameters results: part I. $S_{1,2}$ stands for the monitoring situations, $C_{1,2}$ stands for the subjects groups and p stands for the p-value of Student's t-test among the groups. $S_2 - S_1$ stands for the subtraction of the parameters; it indicates a change in the parameter after the load was applied to the lower limbs. The parameters are: the frequency (f_r) and amplitude (a_r) of the *res-curve*, the amplitude of the *pos-curve* (a_p), the amplitude ratios of the *res-curve* and *pos-curve* (r_{pr}) and the percentage of energy yielded by the three highest spectrum lines (p_3).

		f_r (Hz)	a_r (mm^2)	r_{pr} $(-)$	$a_p(mm^2)$	p_3 $(\%)$
S_1	C_1	0.21 ± 0.06	1823 ± 873	0.3 ± 0.2	378 ± 163	46.7 ± 7.8
	C_2	0.26 ± 0.06	870 ± 297	0.31 ± 0.21	258 ± 152	29.7 ± 6.3
	p	*	**	_	*	**
S_2	C_1	0.22 ± 0.08	1928 ± 864	0.30 ± 0.12	659 ± 353	45.9 ± 5.3
	C_2	0.34 ± 0.1	540 ± 314	0.95 ± 0.61	574 ± 402	24.5 ± 5.8
	p	*	**	*	_	**
$S_2 - S_1$	C_1	0.006 ± 0.07	300 ± 740	-0.02 ± 0.25	160 ± 470	-1.6 ± 6.6
	C_2	0.083 ± 0.08	-360 ± 190	0.7 ± 0.55	320 ± 310	-9.2 ± 3.8
	p	*	*	**	_	*

Table 5.3: The dynamic parameters results: part II. $S_{1,2}$ stands for the monitoring situations, $C_{1,2}$ stands for the subjects groups, p stands for the p-value of Student's t-test among the groups. $S_2 - S_1$ stands for a subtraction of the parameters; it indicates a change in the parameter after the load was applied to the lower limbs. The parameters are: SD (σ_{DC}), the skewness (γ_{DC}) and kurtosis (β_{DC}) of the diff-curve, the range of the diaphragm motion measured at three different points placed on the diaphragm surface ($rg_i, i \in 1, 2, 3$).

		σ_{DC}	γ_{DC}	β_{DC}	rg_1 (mm)	rg_2 (mm)	rg_3 (mm)
S_1	C_1	1416 ± 607	-0.11 ± 0.46	1.92 ± 0.39	21.1 ± 10.1	40.7 ± 13.4	47.1 ± 12.3
	C_2	786 ± 218	-0.65 ± 0.20	2.23 ± 0.33	7 ± 7.7	21.7 ± 5.7	29.8 ± 6.6
	p	**	**	*	**	**	**
S_2	C_1	1711 ± 624	-0.13 ± 0.29	1.67 ± 0.10	22.1 ± 10.8	46.1 ± 14.3	56.5 ± 17.7
	C_2	670 ± 290	-0.57 ± 0.66	2.89 ± 0.68	10.1 ± 6.1	20.6 ± 8.6	23.7 ± 8.1
	p	**	*	**	**	**	**
$S_2 - S_1$	C_1	300 ± 650	0.084 ± 0.37	-0.14 ± 0.51	0.95 ± 11	5.4 ± 15	6.5 ± 19
	C_2	-96 ± 210	0.11 ± 0.66	0.49 ± 0.79	4.6 ± 6.2	-0.36 ± 7	-5.4 ± 8.8
	p	*	–	*	–	–	*

Table 5.4: Static parameters results: $S_{1,2}$ stands for the monitoring situations; $C_{1,2}$ stands for the subjects groups; and p stands for p-value of Student's t-test among the groups. $S_2 - S_1$ stands for a subtraction of the parameters; it indicates a change in the parameter after the load was applied to the lower limbs. The parameters are: diaphragm inclination in the sagittal plane in caudal position (dec_a), the height of a strip overlapping the diaphragm contour (h_d) and the diaphragm height in the thorax (d_p).

		h_d $(-)$	dec_a $(^\circ)$	d_p (mm)
S_1	C_1	0.25 ± 0.06	23.8 ± 7.1	29 ± 28
	C_2	0.32 ± 0.05	15 ± 5.6	-64 ± 18
	p	*	**	**
S_2	C_1	0.25 ± 0.05	24.8 ± 9.6	35 ± 20
	C_2	0.31 ± 0.06	17.8 ± 5.8	-51 ± 17
	p	*	*	**
$S_2 - S_1$	C_1	0.0009 ± 0.04	1.7 ± 6	6.6 ± 20.7
	C_2	-0.02 ± 0.03	3.6 ± 3.1	15.8 ± 14.1
	p	–	–	–

lower limbs (1823 mm^2 S_1, 1928 mm^2 S_2). By contrast, the patients group showed lowered excursions when load was applied (870 mm^2 S_1, 540 mm^2 S_2). The inter-situational difference was significantly different amongst the groups with $p = 0.004$. In comparison with the patients group, the control group had three times bigger excursions in situation S_1, and 6.5 times bigger excursions in situation S_2.

In order to compare diaphragm excursions in millimeters, rg_i parameters were introduced. The diapragm excursion was measured at three points on the diaphragm contour: the anterior, central and posterior parts (see Figure 5.5). The control group exhibited a significantly larger motion range than the patients group in the both situations ($p < 0.001$). In addition, the measurements showed larger motion of the posterior part of the diaphragm than the anterior part. In S_1, the antero-posterior ratio was 2.2 in the control group and 4.2 in the patients group. In S_2, the control group raised the range of the posterior part to 56.5 mm, resulting in the antero-posterior ratio of 2.5. The patients group, in contrast, raised the range in the anterior area and reduced the range in the posterior area, resulting in the antero-posterior ratio of 2.3.

The range of postural movements (the amplitude of the postural curve a_p) was larger in the control group (C_1: 380 mm^2 S_1, 660 mm^2 S_2, C_2: 260 mm^2 S_1, 570 mm^2 S_2), with the only statistically significant difference being in situation S_1 ($p = 0.04$). For both groups, the amplitude of the postural curve rose when a load was applied to the lower limbs, while the rises in C_1 and C_2 did not differ significantly ($p = 0.27$). The amplitude ratio of the *res-curve* and *pos-curve* r_{pr} shows which type of diaphragm motion dominates in the overall motion. When this parameter is greater than 1, the postural moves of the diaphragm are bigger than the respiratory moves, and vice versa. Moreover, in situation S_2 the range of the motion in the patients group was equally distributed between respiratory and postural movement ranges (r_{pr} 0.95, meaning 50 % of the total motion range by the postural motion), while in the both situations the control group had the same r_{pr} of 0.3 (23 % of the total motion range for the postural motion and 77 % for the respiratory motion).

Diaphragm motion harmonicity and central moments

The most important diff-curve shape parameter is its harmonicity, reflected by the parameter p_3. When the patient loses control over diaphragm motion, the diff-curve loses its typical harmonic shape and the three highest spectral lines contain less of the signal energy (see Subsection 5.3.2). The control group was able to keep the harmonicity at almost the same level in both situations (S_1 46.7 %, S_2 46 %), while the patients group achieved a significantly lower ($p < 10^{-7}$) percentage (S_1 29.7 %, S_2 25.5 %). For the patients group, the decrease in the p_3 value was significantly higher ($p = 0.002$) than the decrease for the control group.

The third central statistical moment, skewness (γ_{DC}), elegantly characterizes the centering of the diff-curve around its mean value. This parameter can be used to indicate whether the patient kept the diaphragm longer in the inspiratory position or in the expiratory position. Naturally, harmonic breath would lead to zero skewness. If the diaphragm is kept longer in the inspiratory position, the skewness has a positive value, and if the diaphragm is kept longer in the respiratory position, the skewness has a negative value. In S_1, both the control and the patients group had negative skewness (-0.11,

-0.65). However, the control group exhibited big variance (positive skewness in the case of six subjects). For the patients group, all the values were negative except in the case of one subject. The difference is significant ($p < 0.001$), despite large variance in the control group. In S_2, the mean skewness values were -0.13 and -0.57. The control group became more consistent, while the patients group exhibited large variance in this parameter. This is due to an increase in the influence of the postural part of the diaphragm movement. The difference between the groups C_1 and C_2 was significant ($p = 0.02$). There was no significant change, either in the control or in patients group, after the load was applied to the lower limbs ($p = 0.87$).

The fourth central statistical moment kurtosis can be used to study control over diaphragm movement. The harmonic motion shows lower kurtosis rather than more random, uncontrolled motion. In situation S_1, the control group had a lower kurtosis parameter (1.92) than the patients group (2.23), with a significant difference, $p = 0.03$. In situation S_2, the kurtosis parameter for the control group fell to 1.67, and for the patients group it rose to 2.89, which raised the significance of the inter-group difference (p-value $= 3 \cdot 10^{-6}$).

5.4.2 Static parameters

Diaphragm height, described by the h_d parameter (higher h_d means a more bulging diaphragm), differs significantly between the groups, both in situation S_1 (p-value = 0.001) and in situation S_2 (p-value = 0.003). The parameter was independent of the postural load and has a very similar value for the both situation S_1 and situation S_2: 0.25 for the control group and 0.32 for the patients group. The parameter was normalized by the anteroposterior size of the thorax.

The inclination of the diaphragm in caudal position (dec_a) differs significantly between the groups in the two observed situations (S_1, $p = 0.0005$; S_2, $p = 0.02$). The difference between situation S_1 and situation S_2 was not large (within the SD range) and was statistically the same for the both groups ($p = 0.27$). The mean inclination in situation S_1 was 23.8° in the control group and 15° for the patients group, i.e. the control group kept the diaphragm in a more vertical position.

The diaphragm height in the thorax (d_p) differs considerably between the groups ($p < 10^{-10}$). The control group kept the diaphragm below the back marker in both situations. In situation S_1, the value was 2.9 cm, and in situation S_2, the value was 3.5 cm. The diaphragm was lowered by 0.6 cm on average, which is a small value in comparison with the SD of the values. In situation S_1, the patients group had the diaphragm in a position 6.4 cm above the back marker on average, and 5.1 cm above the marker in situation S_2. The average difference is 1.3 cm. No statistically significant difference (p-value 0.15) was found among the diaphragm shifts after the load was applied to the lower limbs.

5.4.3 Summary

Slower and deeper respiratory motion (parameters f_r, a_r) were concluded in group S_1 in both observed situations. In addition, after the postural demands rose in situation S_2, the breathing speed increased significantly ($p = 0.01$) and the breath depth (a_r) lowered

Figure 5.12: An illustration of the three points selected on the vessel contour for the validation of the motion estimation. Arbitrary points in the anterior and posterior part of the diaphragm contour and the point in the middle of the contour were selected. The arrow shows the average direction of the point movement between successive diaphragm images.

significantly ($p = 0.004$) in the patients group. There were larger postural moves in the control group, and they remained larger in both situations, rising equally for each group. The res/pos ratio, r_{pr}, shows great domination of postural moves in the patients group. As the respiratory movement lowered with the postural load, the ratio increased in the patients group, and the difference between the groups became significant. A significant difference in the harmonicity emerged, which is denoted by the parameters p_3 and β_{DC}. These parameters indicate a much more harmonic diaphragm movement in the control group, with or without the load. The diaphragm motion in the thorax was symmetrical for the control group.

The results for the static parameters revealed that the diaphragm of the control group was flatter (parameter h_d) in both situations. The inclination of the diaphragm was greater (i.e. it was more verticalized) in the control group. The patients group had the diaphragm placed significantly higher in the trunk, as indicated by the d_p parameter.

5.4.4 Diaphragm motion estimation

No reference for the motion estimation was established or provided on the image data available for the presented study. The method of using a diff-curve for motion estimation

Table 5.5: The correlation coefficient of both the vertical and absolute motion of the diaphragm was strongly correlated with the motion as estimated by the diff-area.

	Correlation coefficient					
	Vertical motion			Absolute motion		
	Mean	Min	Max	Mean	Min	Max
p_{ant}	0.968	0.879	0.993	0.967	0.873	0.992
p_{cen}	0.975	0.931	0.997	0.976	0.876	0.995
p_{pos}	0.959	0.863	0.997	0.962	0.855	0.998

has been considered by other research groups [109] and was regarded as the baseline method throughout this work. We compared estimation of the motion by assessing absolute and vertical motion of three points placed equidistantly on the diaphragm contour during the most caudal position of the diaphragm (Figure 5.12). The correlation coefficient has been estimated between the motion of the points and the diff-curve-based motion, the same curves as are illustrated in Figure 5.11.

The resulting correlation coefficient can be found in Table 5.5. The mean correlation coefficient was very high for all the measured points with a mean value over 0.95. This shows strong confidence in the motion estimation. The best correlation was yielded by the motion measured at the central point; this was expected and can be explained by the central point having most stable motion compared to the sides of the diaphragm which are affected more by the rotation of the diaphragm. An important part of diaphragm motion was observed to be in the antero-posterior, as illustrated by the arrow in Figure 5.11.

5.5 Discussion

The extracted parameters were selected in a way that allows a wide spectrum of diaphragm properties to be assessed. Novel methods in the diaphragm motion estimation are the evaluation of the motion harmonicity using statistical methods (skewness, kurtosis), the separation of the diaphragm motion into postural and respiratory parts using the processing of the harmonic spectrum or assessment of the amount of the non-respiratory motion using the energy carried by the spectral lines. A novel method for more precise motion estimation using tracking of the vessel structure visible in the retina has been proposed. From the static parameters, a novel method of diaphragm inclination assessment has been used by fitting a line into the diaphragm contour.

In the results section, we concluded that there is a statistically significant difference in the range of motion (ROM) of the diaphragm. Two and three times greater ROM were noted in the control group compared to the patients group in the both situations S_1 and S_2. In addition, the average diaphragm excursions rg_2 (central part) in situation S_1 were 40 mm in the control group and 22 mm in the patients group. In situation S_2, rg_2 was 46 mm in the control group and 21 mm in the patients group. The diaphragm excursions rose from the ventral part to the dorsal part. Gierada [92] also concluded that there was a bigger motion range in the ventral part of the diaphragm than in the

dorsal part. Kondo, who studied the correlation between lung volume and diaphragm motion, came to the same conclusion in [104]. Kolar [88] observed diaphragm excursions of 27.3 ± 10.2 mm in the apex during tidal breathing and 39 ± 17.6 mm in the dorsal part. Takazakura [105] showed a difference of 20 mm within the highest point of the diaphragm motion when sitting and when supine. Taking into account the large range of diaphragm motions reported in the literature [148], our measurements are consistent.

When considering changes in the range of diaphragm motion after pressure was applied to the lower limbs, the ROM values for the control group rose on the average, but there was large variance in the group, and the rise was bigger in the posterior part than in the anterior part. The ROM values for the patients group rose in the anterior part of the diaphragm, and lessened in the posterior diaphragm part. In contrast to our measurements, Kolar [109] observed the opposite change in the same situations. In Kolar's case, the ROM was the same during tidal breathing, but the group with LBP had lower excursions of the anterior part of the diaphragm. The subjects in Kolar's study had the diaphragm at the same height in the trunk, despite the symptoms. It was observed that the diaphragm was significantly higher for the patients group. This may be a mechanism by which the patients group was able to keep the diaphragm excursions more evenly spread after the postural demands increased.

It was also observed that the diaphragm was more contracted in the posterior part for the control group. Diaphragm inclination measurements showed significant lowering of the posterior part of the diaphragm relative to the anterior part of the diaphragm for the control group. The patients group kept the diaphragm in a more horizontal position. The average changes in inclination after a rise in postural demands were only small in comparison with the variance of the inclination. The height of the diaphragm contour (h_d) above the zone of apposition was also measured as a significant parameter between the groups of subjects. Suwatanapongched [148] concluded that there was a flattening of the diaphragm in the older population in his study. Our results did not show any significant age-related correlation of diaphragm flatness. Instead, the only significant correlation that we observed was between diaphragm height and the LBP intensity of the patients group during the month before the measurements were made. The correlation was significant in situation S_2. It is assumed that this diaphragm bulging is due to a lesser ability to contract the diaphragm properly. To the best of our knowledge, there are no results in the literature for the measurements of diaphragm flatness in subjects suffering from LBP. The lesser ability to contract the diaphragm in the patients group is also supported by the significantly higher position in the trunk.

Another question which emerges in relation to LBP intensity is the effect of acute pain. The effect would bias our findings as the study focused on long-time changes in the motion patterns of the diaphragm. The first factor is the pain induced by the applied load. This was controlled by our methodology and the subjects ensured that no acute pain was induced by postural load. The second factor concerns differences in pain intensity perceived on the day of measurement, and the influence of the pain on the results. An important consideration is that the pain was chronic, and so we assume a tendency of the muscles to overload the spine and some influence on the observed structural degenerative spinal findings. The range of pain intensity on the day of measurement of the patients is wide, ranging from 0 to 8.9. This wide range of pain intensity is useful for revealing a possible dependency of the parameters on the acute perception of pain. The best

practice would be to classify the subjects into groups according to pain and to treat the groups statistically. However, this kind of evaluation was not possible because we would have needed many more study subjects. A second option was to examine the correlation between pain intensity and the measured parameters. No correlation was observed between the measured parameters and pain intensity except for bulging (i.e. long term pain) of the diaphragm, as was discussed above. The results indicate that, as the pain is long term, the patients do not change their respiratory patterns according to fluctuations in the chronic LBP.

It was concluded that a useful method for comparing the ratio of actual respiratory-related motions with other motions of the diaphragm is to separate the differential curve into respiratory and non-respiratory movements of the diaphragm. This division was inspired by various works. In [149, 150, 151] the postural and respiratory functions of the diaphragm were assessed using invasive EMG. Hodges [150, 75, 152] described tonic and respiratory activation during the breathing cycle and superposition of the electromyographic signals phasically related to harmonic limb movement. Hodges also used the harmonic spectrum in his investigation of muscle cooperation for compensating for breathing movements in body posture. Our study showed a non-negligible proportion of non-respiratory diaphragmatic motion, referred to as postural movements. These movements formed one third of the diaphragm motion range, on average, in tidal breathing. The rise in the range of postural motions when there is an increase in postural demands on the body confirms the participation of the diaphragm in postural mechanisms. Separating the respiratory signal from the postural signal was important in cases when postural movements start to form a large proportion of the diaphragm motion, as in situation S_2 for the patients group. A simple investigation of the differential curve does not show significant lowering of the respiratory motion range, but after the signals are separated significant changes are revealed in both the postural and the respiratory parts of the movement.

In our measurements, we did not observe the same diaphragm excursions in the posterior part of the diaphragm for healthy subjects and for subjects suffering from LBP that were observed by [109]. The excursions were reduced in the patients group. In contrast with Kolar's findings [109], we concluded that there was also lowering of the diaphragm's inspiratory position in the patients group in situation S_2. Our measurements support the hypothesis of less diaphragm contraction in the patients group, with a significant correlation between diaphragm bulging and the intensity of the patient's LBP. We did not conclude that any other parameters beyond diaphragm flatness were dependent on the intensity of the subjects' back pain. A high position in the trunk also supports the hypothesis of the lesser ability to contract the diaphragm being found in LBP subjects. These findings support the hypothesis that changed diaphragm recruitment could be an important underlying factor for LBP [11].

An improvement to the state-of-the-art diaphragm processing methods was proposed in the motion detection part. The methods based on the diff-area and diff-curve have been further improved to measure the motion of the diaphragm by registering blood vessels among the successive images. The proposed methods showed themselves to be successful in the registration, correctly assessing the diaphragm motion in all the patients within the group used for testing the motion detection. The enhanced method could have important implications for improving the accuracy of the motion estimation by employing the exact

direction of the motion and could be employed as an alternative method to measuring the height of the diaphragm contraction, as was done in [20]. Another improvement compared to the diff-curve-based method is the direct measurement of the motion in pixels or mm, allowing direct assessment of the motion range of the diaphragm.

Some limitations of the harmonic model of respiratory and postural movements need to be addressed. The modelled breath has to be periodic and preferably harmonious. The breath frequency has to be stable within the observed sequence. If these conditions are not fulfilled the results will be biased. Our measurements were suitable for using the sine model. All subjects displayed a stable frequency of breathing. However it is desirable to extend the model to observe the time dependence of the parameters. The sine wave model of diaphragm postural function works well for assessing the range of postural motion. A more complex model needs to be created for a more detailed inspection of the postural function. Magnetic resonance imaging is a reliable method for making detailed observations and assessments of the diaphragm. A restriction of dynamic assessment is the frequency of the movement. This is limited by the sampling (imaging) frequency, which is currently quite low. Thus the diaphragm can be recruited by only stabilizing compensation in static loadings. A limitation of the automatic processing is the necessity to manually delineate the diaphragm contour in its most caudal and most cranial positions. Besides that, the proposed methodology is independent of the orientation of the patient within the image. Dependency on the resolution of the image exists and has to be taken into account when applying the motion detection method in a different context. The applied blood vessels segmentation method is, however, steerable for images of different resolution.

Chapter VI

Conclusions

In this thesis, diverse features for the characterization of structures in medical images were proposed and validated. The main emphasis was put on the blood vessel segmentation in the retina which was subsequently successfully applied for the vessel segmentation in the diaphragm images. The outcomes described in the thesis are valuable for both research and clinical environments.

6.1 Contributions

The following contributions to the research community were provided in this thesis:

- An overview of the publicly available retinal blood vessels segmentation methods was provided. A quantitative assessment of the methods was made on five publicly available databases while optimizing the method parameters for each database. The methods were compared with the current state-of-the-art. To allow for automatic segmentation of the blood vessels on new databases, a method for predicting the segmentation parameters was proposed. Data from the optimization procedure of the methods were made publicly available[1] and the whole optimization and prediction framework is planned to be released. Results from this part of the work were presented at the Ophthalmic Medical Image Analysis (OMIA) workshop at the Medical Image Computing and Computer Assisted Interventions Conference (MICCAI) 2015 [16] and were published in the journal *Computerized Medical Imaging and Graphics* [17]. A description of the methodology and results is provided in Chapter 3.
- The method for predicting the vessel segmentation parameters was validated through its application on a new database of retinal pictures. It was concluded that the segmentation parameters were estimated successfully and allowed obtaining reliable vessel segmentation. On the basis of vessel segmentation, a new framework

[1] `http://www.it.lut.fi/mvpr/medimg`

for characterizing the retinal vessel structure by automatic estimation of the AVR was proposed. The framework achieved state-of-the-art performance in the vessel classification into arteries and veins. In addition, it contains a novel method for the simple and robust selection of the final vessels used in the estimation of the AVR. An important result was concluded in the observation of a stronger association between the automatically estimated AVR and blood pressure than was observed for the manually estimated AVR. A description of the methodology and results is provided in Chapter 4.

- Software tools were implemented during the work on the retinal characterization. A simple graphical user interface (GUI) for the delineation of the OD in the retina by fitting an ellipse into manually given input points. In addition, a GUI for the labelling of the vessel segments as either arteries or veins and also for labelling the vessel end-points as head or tail. The both tools will be made available to the public for further use.

- A method for the diaphragm motion characterization was proposed in Chapter 5. The method provides an estimate of the proportion of the diaphragm motion which is responsible for respiration, and the proportion which is unrelated to respiration. The features proposed for characterizing the motion can be used to assess the diaphragm motion patterns of patients and make diagnostic decisions for physiotherapic interventions. In addition to the motion characterization, a set of static features characterizing the diaphragm position and shape were proposed and validated on healthy subjects and subjects with back problems. The methods and results from this part were presented at the International Conference on Information Technology and Applications in Biomedicine (ITAB) 2010 [18] and published in the PLOS ONE journal [19].

- A method for the accurate estimation of the diaphragm motion was developed based on registration of the vessel patterns between successive MRI images. The method allows for more accurate detection of the diaphragm movement by tracking arbitrary point on the diaphragm surface. The output of the method is the movement measured in pixels or millimetres. The proposition of the new method and its validation are provided in Chapter 5

6.2 Future work

Several areas exist where the presented research can be continued. The optimization of the vessel segmentation parameters was performed using a grid search. A possibility to apply faster optimization methods can be further investigated. If more intelligent optimization methods would be applicable, it would allow for faster and more easily applicable optimization. The models estimated for the prediction of the segmentation parameters were established using five retinal image databases. This limits the number of predictors that can be reasonably used in the models so including more databases in the process would enable the creation of more general and better models.

During the application of the segmentation methods after the parameter prediction, it was concluded that a methodology allowing proper, application-based assessment of the

resulting vessel segmentation is missing. In the presented case, vessel segmentation was used to establish vessel segments and estimate their edges. The optimal method to assess the quality of automatic vessel segmentation with respect to its following application is a possibility in future research.

The framework for AVR estimation provides many areas for improvement. Among the most important ones is the employment of the vessel structure, i.e., the crossing and bifurcation information, in supporting the classification into arteries and veins. In the future, different classifiers could be considered for the soft classification of the vessel segments. Alternatively, training the classifier using a larger set of images can be considered because, in the current settings, the classifier appeared to fail in some cases where large segments have visually distinct features. Therefore, it is expected that generalizability of the classifier should be improved. An automatic method for OD detection should be added to the framework in order to make it fully autonomous. Finally, the association between the AVR and clinical parameters should take into account diagnoses of the patients.

Future research directions of diaphragm image processing include the development of a fully automated method for ROI detection, i.e., automatic detection of the diaphragm contour. This would be of great interest in order to fully automate the presented method. In addition, the method proposed for registering the diaphragm images is considered as a baseline that justifies the applicability of the vessel structure to be used in the task. However, the method should be compared to the standard registration methods and a faster method should be applied.

6.3 A list of candidate's publications

6.3.1 Publications related to the topic of the thesis

Journals with impact factor

- Vostatek, P., Novak, D., Rychnovský, T., & Rychnovská, Š. (2013). Diaphragm postural function analysis using magnetic resonance imaging. *PLOS ONE, 8(3), e56724.*
  Author's participation: 25 %, *IF(2016)=2.8*

- Hellebrandová, L., Chlumský, J., Vostatek, P., Novák, D., Rýznarová, Z., & Bunc, V. (2016). Airflow limitation is accompanied by diaphragm dysfunction. *Physiological Research, 65(3), 469.*
  Author's participation: 25 %, *IF(2016)=1.5*

- Vostatek, P., Claridge, E., Uusitalo, H., Hauta-Kasari, M., Fält, P., & Lensu, L. (2017). Performance comparison of publicly available retinal blood vessel segmentation methods. *Computerized Medical Imaging and Graphics, 55, 2-12.*
  Author's participation: 50 %, *IF(2016)=1.7*

Other ISI or Scopus indexed publications

- Vostatek, P., Novak, D., Rychnovský, T., & Wild, J. (2010, November). Diaphragm postural function analysis using magnetic resonance. In *Information Technology*

and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on (pp. 1-4). IEEE.
Author's participation: 30 %

Other publications

- Vostatek, P., Claridge, E., Fält, P., Hauta-Kasari, M., Uusitalo, H. & Lensu, L., Evaluation of publicly available blood vessel segmentation methods for retinal images, *Chen X, Garvin MK, Liu JJ, Trusso E, Xu Y, editors. Proceedings of the Ophthalmic Medical Image Analysis Second International Workshop, OMIA 2015, Held in Conjunction with MICCAI 2015, Munich, Germany, October 9, 2015. 137–144.*
Author's participation: 50 %

6.3.2 Publications unrelated to the topic of the thesis

- Macaš, M., Lhotská, L., Bakstein, E., Novák, D., Wild, J., Sieger, T. ... & Jech, R. (2012). Wrapper feature selection for small sample size data driven by complete error estimates. Computer methods and programs in biomedicine, 108(1), 138–150.
Author's participation: 3 %

- Sieger, T., Serranová, T., Růžička, F., Vostatek, P., Wild, J., Šťastná, D. ... & Jech, R. (2015). Distinct populations of neurons respond to emotional valence and arousal in the human subthalamic nucleus. Proceedings of the National Academy of Sciences, 112(10), 3116–3121.
Author's participation: 5 %,

- Bakštein, E., Sieger, T., Wild, J., Nova'k, D., Schneider, J., Vostatek, P. ... & Jech, R. (2017). Methods for Automatic Detection of Artifacts in Microelectrode Recordings. Journal of Neuroscience Methods.
Author's participation: 5 %

[1] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.

[2] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.

[3] R. C. Deo. Machine learning in medicine. *Circulation*, 132(20):1920–1930, 2015.

[4] M. D. Abràmoff, M. K. Garvin, and M. Sonka. Retinal imaging and image analysis. *IEEE transactions on medical imaging*, 3:169–208, 2010.

[5] D. Hoy, C. Bain, G. Williams, L. March, P. Brooks, F. Blyth, A. Woolf, T. Vos, and R. Buchbinder. A systematic review of the global prevalence of low back pain. *Arthritis & Rheumatology*, 64(6):2028–2037, 2012.

[6] J. J. Kanski and B. Bowling. *Synopsis of Clinical Ophthalmology*. Elsevier Health Sciences, 3rd edition, 2013.

[7] G. Liew, J. J. Wang, P. Mitchell, and T. Y. Wong. Retinal vascular imaging a new tool in microvascular disease research. *Circulation: Cardiovascular Imaging*, 1(2):156–161, 2008.

[8] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman. Blood vessel segmentation methodologies in retinal images – a survey. *Comput. Methods Prog. Biomed.*, 108(1):407–433, October 2012.

[9] R Bernardes, P Serranho, and C Lobo. Digital ocular fundus imaging: A review. *Ophthalmologica*, 4:161–81, 2011.

[10] C. Kirbas and F. Quek. A review of vessel extraction techniques and algorithms. *ACM Comput. Surv.*, 36(2):81–121, 2004.

[11] M. D. Smith, A. Russell, and P. W. Hodges. Disorders of breathing and continence have a stronger association with back pain than obesity and physical activity. *The Australian Journal of Physiotherapy*, pages 11–16, 2005.

[12] S. K. Grimstone and P. W. Hodges. Impaired postural compensation for respiration in people with recurrent low back pain. *Experimental Brain Research*, pages 218–224, 2003.

[13] R. Courtney. The function of breathing and its dysfunctions and their relationship to breathing therapy. *International Journal of Osteopathic Medicine*, pages 78–85, 2009.

[14] N. Roussel, J. Nijs, S. Truijen, L. Vervecken, S. Mottram, and G. Stassijns. Altered breathing patterns during lumbopelvic motor control tests in chronic low back pain: a case-control study. *European Spine Journal*, pages 1066–1073, 2009.

[15] L. Janssens, S. Brumagne, K. Polspoel, T. Troosters, and A. McConnell. The effect of inspiratory muscles fatigue on postural control in people with and without recurrent low back pain. *Spine*, pages 1088–1094, 2010.

[16] P. Vostatek, E. Claridge, P. Fält, M. Hauta-Kasari, H. Uusitalo, and L. Lensu. Evaluation of publicly available blood vessel segmentation methods for retinal images. In Chen X, Garvin MK, Liu J, Trucco E, and Xu Y, editors, *Proceedings of the Ophthalmic Medical Image Analysis Second International Workshop, OMIA 2015, Held in Conjunction with MICCAI 2015*, pages 137–144. Iowa Research Online, 2015.

[17] P. Vostatek, E. Claridge, H. Uusitalo, M. Hauta-Kasari, P. Fält, and L. Lensu. Performance comparison of publicly available retinal blood vessel segmentation methods. *Computerized Medical Imaging and Graphics*, 55:2–12, 2017.

[18] P. Vostatek, D. Novák, T. Rychnovský, and J. Wild. Diaphragm postural function analysis using magnetic resonance. In *Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on*, pages 1–4. IEEE, 2010.

[19] P. Vostatek, D. Novák, T. Rychnovský, and Š. Rychnovská. Diaphragm postural function analysis using magnetic resonance imaging. *PLOS ONE*, 8(3):e56724, 2013.

[20] L. Hellebrandová, J. Chlumský, P. Vostatek, D. Novák, Z. Rýznarová, and V. Bunc. Airflow limitation is accompanied by diaphragm dysfunction. *Physiological Research*, 65(3):469, 2016.

[21] H. Kolb. Gross anatomy of the eye. 2007.

[22] Blausen.com staff. Medical gallery of blausen medical 2014. WikiJournal of Medicine 1 (2), DOI: 10.15347/wjm/2014.010, ISSN: 2002-4436.

[23] K. Rothaus, X. Jiang, and P. Rhiem. Separation of the retinal vascular graph in arteries and veins based upon structural knowledge. *Image and Vision Computing*, 27(7):864–875, 2009.

[24] F. Carinci, F. Stracci, and F. Storms. Diabetes. In *Major and Chronic Diseases, Report 2007*. European Union, Directorate–General for Health and Consumers, 2007. ISBN 92-79-08896-4.

[25] Jack J Kanski and Brad Bowling. *Clinical ophthalmology: a systematic approach.* Elsevier Health Sciences, 2011.

[26] N. Patton, T. M. Aslam, T. MacGillivray, I. J. Deary, B. D., R. H. Eikelboom, K. Yogesan, and I. J. Constabl. Retinal image analysis: Concepts, applications and potential. *Progress in Retinal and Eye Research*, 25(1):99 –127, January 2005.

[27] Cong Sun, Jie Jin Wang, David A Mackey, and Tien Y Wong. Retinal vascular caliber: systemic, environmental, and genetic associations. *Survey of ophthalmology*, 54(1):74–95, 2009.

[28] N. Chapman, G. Dell'Omo, M. S. Sartini, N. Witt, A. Hughes, S. Thom, and R. Pedrinelli. Peripheral vascular disease is associated with abnormal arteriolar diameter relationships at bifurcations in the human retina. *Clinical Science*, 103(2):111–116, 2002.

[29] O. Brinchmann-Hansen. The light reflex on retinal arteries and veins. a theoretical study and a new technique for measuring width and intensity profiles across retinal vessels. *Acta Ophthalmologica. Supplement*, 179:1–53, 1985.

[30] J. C. Parr and G. F. S. Spears. General caliber of the retinal arteries expressed as the equivalent width of the central retinal artery. *American Journal of Ophthalmology*, 77(4):472–477, 1974.

[31] L. D. Hubbard, B. Ehrhardt, R. Klein, S. Messing, R. J. Brothers, S. E. Moss, and S. M. Meuer. The association between generalized retinal arteriolar narrowing and blood-pressure. In *Investigative Ophthalmology & Visual Science*, volume 33, pages 804–804. Lippincott-Raven Publ. 227 East Washington SQ, Philadelphia, PA 19106, 1992.

[32] L. D. Hubbard, R. J. Brothers, W. N. King, L. X. Clegg, R. Klein, Lawton S. C., A. R. Sharrett, M. D. Davis, and J. Cai. Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study. *Ophthalmology*, 106(12):2269–2280, 1999.

[33] M. D. Knudtson, K. E. Lee, L. D. Hubbard, T. Y. Wong, R. Klein, and B. E. K. Klein. Revised formulas for summarizing retinal vessel diameters. *Current Eye Research*, 27(3):143–149, 2003.

[34] M. K. Ikram, F. J. de Jong, J. R. Vingerling, J. C. M. Witteman, A. Hofman, M. M. B. Breteler, and P. T. V. M. de Jong. Are retinal arteriolar or venular diameters associated with markers for cardiovascular disorders? the rotterdam study. *Investigative Ophthalmology & Visual Science*, 45(7):2129–2134, 2004.

[35] T. Y. Wong, B. B. Duncan, S. H. Golden, R. Klein, D. J. Couper, B. E. K. Klein, L. D. Hubbard, A. R. Sharrett, and M. I. Schmidt. Associations between the metabolic syndrome and retinal microvascular signs: the atherosclerosis risk in communities study. *Investigative Ophthalmology & Visual Science*, 45(9):2949–2954, 2004.

[36] T. Y. Wong, R. Klein, A. R. Sharrett, B. B. Duncan, D. J. Couper, B. E. K. Klein, L. D. Hubbard, and F. J. Nieto. Retinal arteriolar diameter and risk for hypertension. *Annals of Internal Medicine*, 140(4):248–255, 2004.

[37] J. Soares, J. Leandro, R. Cesar, H. Jelink, and M. Cree. Retinal vessel segmentation using the 2d gabor wavelet and supervised classification. *IEEE Transactions on Medical Imaging*, 25(9):1214–1222, September 2006.

[38] M. Sofka and C. V. Stewart. Retinal vessel extraction using multiscale matched filters, confidence and edge measures. *IEEE Transactions on Medical Imaging*, 25(12):1531–1546, December 2006.

[39] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov. Trainable cosfire filters for vessel delineation with application to retinal images. *Medical Image Analysis*, 19(1):46–57, 2015.

[40] P. Bankhead, C. N. Scholfield, J. G. McGeown, and T. M. Curtis. Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PLOS ONE*, 7(3):e32435, 03 2012.

[41] J. I. Starck, J. Fadili, and F. Murtagh. The undecimated wavelet decomposition and its reconstruction. *IEEE Transactions on Image Processing*, 16:297–309, 2007.

[42] U. T. V. Nguyen, Al. Bhuiyan, L. A. F. Park, and K. Ramamohanarao. An effective retinal blood vessel segmentation method using multi-scale line detection. *Pattern Recognition*, 46(3):703–715, 2013.

[43] E. Ricci and R. Perfetti. Retinal blood vessel segmentation using line operators and support vector classification. *IEEE Transactions on Medical Imaging*, 26(10):1357–1365, 2007.

[44] D. Relan, T. MacGillivray, L. Ballerini, and E. Trucco. Automatic retinal vessel classification using a least square-support vector machine in vampire. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 142–145. IEEE, 2014.

[45] D. Relan, L. Ballerini, E. Trucco, and T. MacGillivray. Retinal vessel classification based on maximization of squared-loss mutual information. In *Machine Intelligence and Signal Processing*, pages 77–84. Springer, 2016.

[46] R. Chrástek, M. Wolf, K. Donath, H. Niemann, D. Paulus, T. Hothorn, B. Lausen, R. Lämmer, C. Y. Mardin, and G. Michelson. Automated segmentation of the optic nerve head for diagnosis of glaucoma. *Medical Image Analysis*, 9(4):297–314, 2005.

[47] J. A. K. Suykens, T. Van Gestel, and J. De Brabanter. *Least squares support vector machines*. World Scientific, 2002.

[48] M. Sugiyama, G. Niu, M. Yamada, M. Kimura, and H. Hachiya. Information-maximization clustering based on squared-loss mutual information. *Neural Computation*, 26(1):84–131, 2014.

[49] E. Grisan and A. Ruggeri. A divide et impera strategy for automatic classification of retinal vessels into arteries and veins. In *Engineering in medicine and biology society, 2003. Proceedings of the 25th annual international conference of the IEEE*, volume 1, pages 890–893. IEEE, 2003.

[50] James C Bezdek. *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media, 2013.

[51] E. Grisan, M. Foracchia, and A. Ruggeri. Color fundus images luminance and contrast normalization. In *3rd CAFIA workshop, in European Journal of Ophthalmology*, pages 228–229, 2003.

[52] L. Tramontan, E. Grisan, and A. Ruggeri. An improved system for the automatic estimation of the arteriolar-to-venular diameter ratio (avr) in retinal images. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 3550–3553. IEEE, 2008.

[53] C. Kondermann, D. Kondermann, M. Yan, et al. Blood vessel classification into arteries and veins in retinal images. In *Proceedings of SPIE Medical Imaging*, pages 651247–6512479, 2007.

[54] C. Nieuwenhuis and M. Yan. Knowledge based image enhancement using neural networks. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 3, pages 814–817. IEEE, 2006.

[55] M. Saez, S. González-Vázquez, M. González-Penedo, M. A. Barceló, M. Pena-Seijo, G. Coll de Tuero, and A. Pose-Reino. Development of an automated system to classify retinal vessels into arteries and veins. *Computer Methods and Programs in Biomedicine*, 108(1):367–376, 2012.

[56] M. Niemeijer, B. van Ginneken, M. D. Abràmoff, et al. Automatic classification of retinal vessels into arteries and veins. *Medical imaging*, pages 72601F–72601F, 2009.

[57] M. Niemeijer, X. Xu, A. V. Dumitrescu, P. Gupta, B. van Ginneken, J. C. Folk, and M. D. Abramoff. Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs. *IEEE Transactions On Medical Imaging*, 30(11):1941–1950, 2011.

[58] C. Muramatsu, Y. Hatanaka, T. Iwase, T. Hara, and H. Fujita. Automated selection of major arteries and veins for measurement of arteriolar-to-venular diameter ratio on retinal fundus images. *Computerized Medical Imaging and Graphics*, 35(6):472–480, 2011.

[59] B. Dashtbozorg, A. M. Mendonça, and A. Campilho. An automatic graph-based approach for artery/vein classification in retinal images. *IEEE Transactions on Image Processing*, 23(3):1073–1083, 2014.

[60] B. Dashtbozorg, A. M. Mendonça, and A. Campilho. Assessment of retinal vascular changes through arteriolar-to-venular ratio calculation. In *International Conference Image Analysis and Recognition*, pages 335–343. Springer, 2015.

[61] M. Foracchia, E. Grisan, and A. Ruggeri. Luminosity and contrast normalization in retinal images. *Medical Image Analysis*, 9(3):179–190, June 2005.

[62] J. Fairfield. Toboggan contrast enhancement for contrast segmentation. In *Pattern Recognition, 1990. Proceedings., 10th International Conference on*, volume 1, pages 712–716. IEEE, 1990.

[63] X. Yao and Y. Hung. Fast image segmentation by sliding in the derivative terrain. In *Intelligent Robots and Computer Vision X: Algorithms and Techniques*, volume 1607, pages 369–379, 1992.

[64] D. J. J. Farnell, F. N. Hatfield, P. Knox, M. Reakes, S. Spencer, D. Parry, and S. P. Harding. Enhancement of blood vessels in digital fundus photographs via the application of multiscale line operators. *Journal of the Franklin Institute*, 345(7):748–765, 2008.

[65] C. G. Owen, A. R. Rudnicka, R. Mullen, S. A. Barman, D. Monekosso, P. H. Whincup, J. Ng, and C. Paterson. Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (caiar) program. *Investigative Ophthalmology & Visual Science*, 50(5):2004–2010, 2009.

[66] J. J. Staal, M. D. Abrámoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, April 2004.

[67] J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou. Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. *IET Image Processing*, 7(4):373–383, 2013.

[68] A. Hoover, V. Kouznetsova, and M. Goldbaum. Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210, March 2000.

[69] S. G. Vázquez, N. Barreira, M. G. Penedo, M. Ortega, and A. Pose-Reino. Improvements in retinal vessel clustering techniques: towards the automatic computation of the arterio venous ratio. *Computing*, 90(3-4):197–217, 2010.

[70] R. Estrada, C. Tomasi, S. C. Schmidler, and S. Farsiu. Tree topology estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(8):1688–1701, 2015.

[71] B. Bordoni and E. Zanier. Anatomic connections of the diaphragm: Influence of respiration on the body system. *Journal of Multidisciplinary Healthcare*, 6:281, 2013.

[72] Wikipedia: Thoracic diaphragm. https://en.wikipedia.org/wiki/Thoracic_diaphragm. Accessed: 2017-08-16.

[73] K. Lewit. Relation of faulty respiration to posture and with clinical implications. *The Journal of the American Osteopathic Association*, pages 525–529, 1980.

[74] K. P. Barr, M. Griggs, and T. Cadby. Lumbar stabilization: Core concepts and current literature, part 1. *American Journal of Physical Medicine and Rehabilitation*, pages 473–480, 2005.

[75] P. W. Hodges and C. A. Richardson. Altered trunk muscle recruitment in people with low back pain with upper limb movement at different speeds. *Archives of Physical Medicine and Rehabilitation*, pages 1005–1012, 1999.

[76] J. Cholewicki, P. C. Ivancic, and A. Radebold. Can increased intra-abdominal pressure in humans be decoupled from trunk muscle co-contraction during steady state isometric exertions? european journal of applied physiology. *European journal of applied physiology*, pages 127–133, 2002.

[77] E. Cohen, A. Mier, and P. Haywood et al. Excursion-volume relation of the right hemidiaphragm measured by ultrasonography and respiratory airflow measurement. *Thorax*, pages 885–889, 1994.

[78] I. A. Kapandji. *The physiology of the joints. Volume 3. The trunk and the vertebral column.* Churchill Livingstone, London, 1998.

[79] A. De Troyer and M. Estenne. Limitations of measurement of transdiaphragmatic pressure in detecting diaphragmatic weakness. *Thorax*, pages 169–174, 1981.

[80] J. Cholewicki and S. M. McGill. Mechanical stability of the in viva lumbar spine: Implications for injury and chronic low back pain. *Clinical biomechanics (Bristol, Avon)*, pages 1–15, 1996.

[81] T. Iwasawa, S. Kagei, T. Gotoh, and et al. Magnetic resonance analysis of abnormal diaphragmatic motion in patients with emphysema. *European Respiratory Journal*, pages 225–231, 2002.

[82] P. W. Hodges, J. E. Butler, D. K. McKenzie, and et al. Contraction of the human diaphragm during rapid postural adjustments. *Journal of Physiology*, pages 539–548, 1997.

[83] P. W. Hodges, C. A. Richardson, and et al. Relationship between limb movement speed and associated contraction of the trunk muscles. *Ergonomics*, pages 539–548, 1997.

[84] P. W. Hodges, S. C. Gandevia, and C. A. Richardson et al. Contractions of specific abdominal muscles in postural tasks are affected by respiratory maneuvers. *Journal of Applied Physiology*, pages 753–760, 1997.

[85] P. W. Hodges. Changes in motor planning of feedforward postural responses of the trunk muscles in low back pain. *Experimental Brain Research.*, pages 261–266, 2001.

[86] J. T. Sharp, N. B. Goldberg, W. S. Druz, H. C. Fishman, and J. Danon. Thoracoabdominal motion in chronic obstructive pulmonary disease. *The American Review of Respiratory Disease.*, pages 47–56, 1977.

[87] A. De Troyer. Effect of hyperinflation on the diaphragm. *The European Respiratory Journal: Official Journal of the European Society for Clinical Respiratory Physiology.*, pages 708–713, 1997.

[88] P. Kolar, J. Neurwith, J. Sanda, V. Suchanek, Z. Svata, J. Volejnik, and M. Pivec. Analysis of diaphragm movement during tidal breathing and during its activation while breath holding using mri synchronized with spirometry. *Physiol Res*, (3):383–392, 2009.

[89] M. C. Jensen, M. N. Brant-Zawadzki, N. Obuchowski, M. T. Modic, D. Malkasian, and J. S. Ross. Magnetic resonance imaging of the lumbar spine in people without back pain. *The New England Journal of Medicine.*, pages 69–73, 1994.

[90] E. Carragee, T. Alamin, I. Cheng, T. Franklin E. van den Haak, and E. Hurwitz. Are first-time episodes of serious lbp associated with new mri findings? *Spine J.*, pages 624–635, 2006.

[91] S. M. McGill, M. T. Sharratt, and J. P. Seguin. Loads on spinal tissues during simultaneous lifting and ventilatory challenge. *Ergonimics*, pages 1772–1792, 1995.

[92] D. Gierada, J. Curtin, S.Erickson, R. Prost, J. Strandt, and L. Goodman. Diaphragmatic motion: Fast gradient-recalled-echo mr imaging in healthy subjects. *Radiology*, pages 879–884, 1995.

[93] D. S. Gierada, J. J. Curtin, and S. J. Erickson et al. Fast gradient echo magnetic resonance imaging of the normal diaphragm. *Journal of Thoracic Imaging*, pages 70–74, 1997.

[94] C. Plathow, H. Zimmermann, and C. Fink et al. Influence of different breathing maneuvers on internal and external organ motion: use of fiducial markers in dynamic mri. *International Journal of Radiation Oncology, Biology, Physics*, pages 238–245, 2005.

[95] C. Plathow, M. Schoebinger, C. Fink, and S. Ley. Evaluation of lung volumetry using dynamic three-dimensional magnetic resonance imaging. *Investigative Radiology*, pages 173–179, 2005.

[96] C. Plathow, C. Fink, A. Sander, and et al. Comparison of relative forced expiratory volume of one second with dynamic magnetic resonance imaging parameters in healthy subjects and patients with lung cancer. *Journal of Magnetic Resonance Imaging*, pages 212–218, 2005.

[97] C. Plathow, S. Ley, and J. Zaporozhan et al. Assessment of reproducibility and stability of different breath-hold maneuvres by dynamic mri: comparison between healthy adults and patients with pulmonary hypertension. *European Radiology*, pages 173–179, 2006.

[98] K. Suga, T. Tsukuda, and H. Awaya. Impaired respiratory mechanics in pulmonary emphysema: evaluation with dynamic breathing mri. *J Magn Reson Imaging*, pages 510–520, 1999.

[99] T. Kotani, S. Minami, and K. Takahashi et al. An analysis of chest wall and diaphragm motions in patients with idiopathic scoliosis using dynamic breathing mri. *Spine*, pages 298–302, 2004.

[100] W. C. Chu and A. M. Li. Dynamic magnetic resonance imaging in assessing lung function in adolescent idiopathic scoliosis: a pilot study of comparison before and after posterior spinal fusion. *Spine*, pages 2243–2249, 2006.

[101] K. Daggfeldt and A. Thorstensson. The mechanics of back-extensor torque production about the lumbar spine. *Journal of Biomechanics*, pages 815–825, 2003.

[102] O. Etlik, M. E. Sakarya, K. Uzun, M. Harman, O Temizoz, and A. Durmus. Demonstrating the effect of theophylline treatment on diaphragmatic movement in chronic obstructive pulmonary disease patients by mrfluoroscopy. *Eur J Radiol*, pages 150–154, 2004.

[103] Ö. Ünal, H. Arslan, K. Uzun, B. Ozbay, and M. E. Sakarya. Evaluation of diaphragmatic movement with mr fluoroscopy in chronic obstructive pulmonary disease. *Clin Imaging*, pages 347–350, 2000.

[104] T. Kondo, I. Kobayashi, and Y. Taguchi et al. A dynamic analysis of chest wall motions with mri in healthy young subjects. *Respirology*, pages 19–25, 2000.

[105] R. Takazakura, M. Takahashi, N. Nitta, and K. Murata. Diaphragmatic motion in the sitting and supine positions: Healthy subject study using a vertically open magnetic resonance system. *J Magn Reson Imaging*, (5):605–609, 2004.

[106] C. Plathow, C. Fink, and S. Ley. Measurement of diaphragmatic length during the breathing cycle by dynamic mri: comparison between healthy adults and patients with an intrathoracic tumor. european radiology. *Eur Radiol*, pages 1392–1399, 2004.

[107] K. Miyamoto, S. Katsuji, and M. Kazuaki. Fast mri used to evaluate the effect of abdominal belts during contraction of trunk muscles. *Spine*, pages 1749–1754, 2002.

[108] B. Hungerford, W. Gilleard, and P. W. Hodges. Evidence of altered lumbopelvic muscle recruitment in the presence of sacroiliac joint pain. *Spine*, pages 1593–1600, 2003.

[109] P. Kolar, J. Sulc, M. Kyncl, J. Sanda, O. Cakrt, R. Andel, K. Kumagai, and A. Kobesova. Postural function of the diaphragm in persons with and without chronic low back pain. *The Journal of Orthopaedic and Sports Physical Therapy*, (4):352–362, 2012.

[110] Muhammad Moazam Fraz, Paolo Remagnino, Andreas Hoppe, Bunyarit Uyyanonvara, Alicja R Rudnicka, Christopher G Owen, Sarah Barman, et al. An ensemble classification-based approach applied to retinal blood vessel segmentation. *Biomedical Engineering, IEEE Transactions on*, 59(9):2538–2548, 2012.

[111] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi. Blood vessel segmentation of fundus images by major vessel extraction and subimage classification. *IEEE Journal of Biomedical and Health Informatics*, 19(3):1118–1128, 2015.

[112] J. I. Orlando and M. Blaschko. Learning fully-connected crfs for blood vessel segmentation in retinal images. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 634–641. Springer, 2014.

[113] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing*, 149:708–717, 2015.

[114] Elahe Moghimirad, Seyed Hamid Rezatofighi, and Hamid Soltanian-Zadeh. Retinal vessel segmentation using a multi-scale medialness function. *Computers in biology and medicine*, 42(1):50–60, 2012.

[115] Zhiyong Xiao, Mouloud Adel, and Salah Bourennane. Bayesian method with spatial constraint for retinal vessel segmentation. *Computational and mathematical methods in medicine*, 2013, 2013.

[116] E. Imani, M. Javidi, and H. Pourreza. Improvement of retinal blood vessel detection using morphological component analysis. *Computer Methods and Programs in Biomedicine*, 118(3):263–279, 2015.

[117] L. Zhang, M. Fisher, and W. Wang. Retinal vessel segmentation using multi-scale textons derived from keypoints. *Computerized Medical Imaging and Graphics*, 45:47–56, 2015.

[118] Y. Q. Zhao, X. H. Wang, X. F. Wang, and F. Y. Shih. Retinal vessels segmentation based on level set and region growing. *Pattern Recognition*, 47(7):2437–2446, 2014.

[119] M. Krause, R. M. Alles, B. Burgeth, and J. Weickert. Fast retinal vessel analysis. *Journal of Real-Time Image Processing*, pages 1–10, 2013.

[120] X. Liu, Z. Zeng, and X. Wang. Vessel segmentation in retinal images with a multiple kernel learning based method. In *Neural Networks (IJCNN), 2014 International Joint Conference on*, pages 507–511. IEEE, 2014.

[121] N. Strisciuglio, G. Azzopardi, M. Vento, and N. Petkov. Multiscale blood vessel delineation using b-cosfire filters. In *International Conference on Computer Analysis of Images and Patterns*, pages 300–312. Springer, 2015.

[122] I. Lázár and A. Hajdu. Segmentation of retinal vessels by means of directional response vector similarity and region growing. *Computers in Biology and Medicine*, 66:209–221, 2015.

[123] A. Mendonça, B. Dashtbozorg, and A. Campilho. Segmentation of the vascular network of the retina. *Image Analysis and Modeling in Opthalmology*, pages 85–109, 2014.

[124] M. S. Miri and A. Mahloojifar. Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction. *Biomedical Engineering, IEEE Transactions on*, 58(5):1183–1192, 2011.

[125] M. M. Fraz, P. Remagnino, A. Hoppe, and S. A. Barman. Retinal image analysis aimed at extraction of vascular structure using linear discriminant classifier. In *Computer Medical Applications (ICCMA), 2013 International Conference on*, pages 1–6. IEEE, 2013.

[126] J. Zhang, E. Bekkers, S. Abbasi, B. Dashtbozorg, and B. ter Haar Romeny. Robust and fast vessel segmentation via gaussian derivatives in orientation scores. In *International Conference on Image Analysis and Processing*, pages 537–547. Springer, 2015.

[127] B. Perret and C. Collet. Connected image processing with multivariate attributes: an unsupervised markovian classification approach. *Computer Vision and Image Understanding*, 133:1–14, 2015.

[128] X. You, Q. Peng, Y. Yuan, Y. Cheung, and J. Lei. Segmentation of retinal blood vessels using the radial projection and semi-supervised approach. *Pattern Recognition*, 44(10):2314–2324, 2011.

[129] B. Yin, H. Li, B. Sheng, X. Hou, Y. Chen, W. Wu, P. Li, R. Shen, Y. Bao, and W. Jia. Vessel extraction from non-fluorescein fundus images using orientation-aware detector. *Medical Image Analysis*, 26(1):232–242, 2015.

[130] F. Argüello, D. L. Vilariño, D. B. Heras, and A. Nieto. Gpu-based segmentation of retinal blood vessels. *Journal of Real-Time Image Processing*, pages 1–10, 2014.

[131] R. Masooomi, A. Ahmadifard, and A. Mohtadizadeh. Retinal vessel segmentation using non-subsampled contourlet transform and multi-scale line detection. In *Intelligent Systems (ICIS), 2014 Iranian Conference on*, pages 1–5. IEEE, 2014.

[132] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 130–137. Springer, 1998.

[133] D. Kaba, C. Wang, Y. Li, A. Salazar-Gonzalez, X. Liu, and A. Serag. Retinal blood vessels extraction using probabilistic modelling. *Health Information Science and Systems*, 2(1):2, 2014.

[134] B. Zhang, L. Zhang, L. Zhang, and F. Karray. Retinal vessel extraction by matched filter with first-order derivative of gaussian. *Comput. Biol. Med.*, 40(4):438–445, 2010.

[135] Q. Li, J. You, and D. Zhang. Vessel segmentation and width estimation in retinal images using multiscale production of matched filter responses. *Expert Systems with Applications*, 39(9):7600–7610, 2012.

[136] R. Annunziata, A. Garzelli, L. Ballerini, A. Mecocci, and E. Trucco. Leveraging multiscale hessian-based enhancement with a novel exudate inpainting technique for retinal vessel segmentation. *IEEE Journal of Biomedical and Health Informatics*, 20(4):1129–1138, 2016.

[137] A. Christodoulidis, T. Hurtut, H. B. Tahar, and F. Cheriet. A multi-scale tensor voting approach for small retinal vessel segmentation in high resolution fundus images. *Computerized Medical Imaging and Graphics*, 2016.

[138] M. Al-Rawi, M. Qutaishat, and M. Arrar. An improved matched filter for blood vessel detection of digital retinal images. *Computers in Biology and Medicine*, 37(2):262–267, 2007.

[139] B. S. Y. Lam, Y. Gao, and A. W. Liew. General retinal vessel segmentation using regularization-based multiconcavity modeling. *Medical Imaging, IEEE Transactions on*, 29(7):1369–1381, 2010.

[140] M. R. N. Tagore, G. B. Kande, E. V. K. Rao, and B. P. Rao. Segmentation of retinal vasculature using phase congruency and hierarchical clustering. In *Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on*, pages 361–366. IEEE, 2013.

[141] E. Cheng, L. Du, Y. Wu, Y. J. Zhu, V. Megalooikonomou, and H. Ling. Discriminative vessel segmentation in retinal images by fusing context-aware hybrid features. *Machine Vision and Applications*, 25(7):1779–1792, 2014.

[142] E. Trucco, A. Ruggeri, T. Karnowski, L. Giancarlo, E. Chaum, J. P. Hubschman, B. Al-Diri, C. Y. Cheung, D. Wong, M. Abramoff, G. Lim, D. Kumar, P. Burlina, N. M. Bressler, H. Jelinek, F. Maiaudeau, G. Quellec, T. MacGillivray, and B. Dhillon. Validating retinal fundus image analysis algorithms: Issues and a proposal. *Investigative Ophthalmology & Visual Science*, pages 3546–3559, 2013.

[143] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on Optimization*, 9(1):112–147, 1998.

[144] T. Y. Wong, M. D. Knudtson, R. Klein, B. E. K. Klein, S. M. Meuer, and L. D. Hubbard. Computer-assisted measurement of retinal vessel diameters in the beaver dam eye study: methodology, correlation between eyes, and effect of refractive errors. *Ophthalmology*, 111(6):1183–1190, 2004.

[145] U. T. V. Nguyen, K. Ramamohanarao, L. A. F. Park, L. Wang, and A. Bhuiyan. A quantitative measure for retinal blood vessel segmentation evaluation. *Int. J. Comput. Vis. Signal Process*, 1(1):1–8, 2012.

[146] Y. Gavet, M. Fernandes, J. Debayle, and J. Pinoli. Dissimilarity criteria and their comparison for quantitative evaluation of image segmentation: application to human retina vessels. *Machine vision and applications*, 25(8):1953–1966, 2014.

[147] V. Janda. *Funkční svalový test [in czech]*. Grada Publishing, Praha, 1996.

[148] T. Suwatanapongched, D. S. Gierada, R. M. Slone, T. K. Pilgram, and P. G. Tuteur. Variation in diaphragm position and shape in adults with normal pulmonary function. *Chest*, pages 2019–2027, 2003.

[149] J. Skladal and C. Ruth. Some remarks concerning the postural activity of the diaphragm in man. *Agressologie*, pages 45–46, 1978.

[150] P. W. Hodges and S. C. Gandevia. Activation of the human diaphragm during a repetitive postural task. *The Journal of Physiology*, pages 165–175, 2000.

[151] J. E. Butler, D. K. McKenzie, and S. C. Gandevia. Discharge frequencies of single motor units in human diaphragm and parasternal muscles in lying and standing. *J Appl Physiol*, pages 147–154, 2001.

[152] P. W. Hodges, V. S. Gurfinkel, S. Brumagne, T. C. Smith, and P. C. Cordo. Coexistence of stability and mobility in postural control: evidence from postural compensation for respiration. *Experimental Brain Research*, pages 293–302, 2002.

Appendix I

Retinal blood vessel segmentation

I.1 Models of the non resolution-related parameters

The models of the non-resolution related parameters of the methods for the retinal blood vessel segmentation. Properties of the models are described in Section 3.3.2.

Table I.1: *Azzopardi method*, σ_{01}, σ_{02}

Predictors (Acc)	R^2 median (min, max)	Predictors (AUC)	R^2 median (min, max)
d_{v3}	0.13 (-0.27, 0.53)	d_{v2}	0.62 (0.47, 0.76)
d_{v1}	0.12 (-0.19, 0.43)	d_{v1}	0.61 (0.59, 0.62)
d_r	0.11 (-0.13, 0.35)	d_{v3}	0.58 (0.48, 0.69)
d_{v2}	0.10 (-0.31, 0.52)	d_r	0.46 (0.45, 0.47)
d_{od}	0.06 (-0.12, 0.25)	d_{od}	0.44 (0.43, 0.45)
d_{v2}, d_{v3}	0.29 (0.29, 0.29)	d_r, d_{v1}	0.72 (0.65, 0.80)
d_{df}, d_{nf}	0.13 (-0.54, 0.80)	d_{v1}, d_{od}	0.58 (0.49, 0.66)
d_{v1}, d_{v2}	0.11 (-0.06, 0.28)	d_{v1}, d_{v2}	0.52 (0.39, 0.65)
d_r, d_{v2}	0.10 (-0.07, 0.28)	d_{df}, d_{nf}	0.50 (0.30, 0.71)
d_r, d_{v3}	0.07 (-0.20, 0.35)	d_{v2}, d_{df}	0.50 (0.34, 0.65)
d_r, d_{od}, d_{nf}	0.87 (0.74, 1.00)	d_r, d_{v1}, d_{v2}	0.77 (0.61, 0.93)
d_{v1}, d_{v2}, d_{df}	0.75 (0.58, 0.93)	d_{v2}, d_{od}, d_{df}	0.73 (0.69, 0.77)
d_r, d_{od}, d_{df}	0.64 (0.31, 0.97)	d_r, d_{v1}, d_{v3}	0.73 (0.60, 0.86)
d_{v1}, d_{v2}, d_{nf}	0.63 (0.40, 0.85)	d_{v3}, d_{od}, d_{df}	0.69 (0.42, 0.95)
d_r, d_{v2}, d_{v3}	0.24 (0.06, 0.42)	d_r, d_{v1}, d_{df}	0.68 (0.60, 0.76)

Table I.2: *Azzopardi method*, a_1, a_2

Predictors (Acc)	R^2 median (min, max)	Predictors (AUC)	R^2 median (min, max)
d_{v2}	0.65 (0.32, 0.97)	d_{nf}	-0.09 (-0.25, 0.06)
d_{v3}	0.55 (0.18, 0.92)	d_{df}	-0.10 (-0.26, 0.05)
d_{v1}	0.37 (-0.03, 0.77)	d_{od}	-0.16 (-0.17, -0.15)
d_r	0.25 (-0.16, 0.66)	d_r	-0.18 (-0.20, -0.16)
d_{od}	0.17 (-0.22, 0.57)	d_{v1}	-0.22 (-0.22, -0.22)
d_r, d_{v2}	0.95 (0.93, 0.98)	d_{od}, d_{df}	0.09 (-0.34, 0.51)
d_{v2}, d_{od}	0.95 (0.92, 0.98)	d_{v1}, d_{v3}	0.00 (-0.55, 0.56)
d_r, d_{v3}	0.95 (0.92, 0.97)	d_r, d_{df}	-0.07 (-0.49, 0.34)
d_{v1}, d_{od}	0.91 (0.86, 0.97)	d_{v1}, d_{v2}	-0.12 (-0.72, 0.48)
d_{v3}, d_{od}	0.91 (0.87, 0.96)	d_{od}, d_{nf}	-0.18 (-0.58, 0.23)
d_{v1}, d_{v2}, d_{od}	0.99 (0.99, 1.00)	d_{v1}, d_{v2}, d_{od}	0.97 (0.96, 0.99)
d_{v2}, d_{od}, d_{df}	0.97 (0.97, 0.97)	d_{od}, d_{df}, d_{nf}	0.87 (0.76, 0.97)
d_r, d_{v1}, d_{v3}	0.96 (0.94, 0.98)	d_{v3}, d_{df}, d_{nf}	0.82 (0.66, 0.99)
d_{v3}, d_{od}, d_{nf}	0.96 (0.93, 0.99)	d_{v1}, d_{v3}, d_{nf}	0.81 (0.64, 0.98)
d_{v3}, d_{od}, d_{df}	0.96 (0.92, 0.99)	d_r, d_{df}, d_{nf}	0.79 (0.57, 1.00)

Table I.3: *Azzopardi method*, τ

Predictors (Acc)	R^2 median (min, max)	Predictors (AUC)	R^2 median (min, max)
d_{v2}	0.05	d_{od}	0.58
d_{v3}	0.03	d_{nf}	0.57
d_r	-0.02	d_r	0.57
d_{od}	-0.03	d_{df}	0.57
d_{v1}	-0.05	d_{v1}	0.39
d_{od}, d_{df}	-0.20	d_{v2}, d_{v3}	0.91
d_{v1}, d_{v3}	-0.34	d_r, d_{v1}	0.88
d_{v1}, d_{v2}	-0.37	d_r, d_{v2}	0.69
d_r, d_{df}	-0.38	d_r, d_{v3}	0.63
d_{od}, d_{nf}	-0.39	d_{v1}, d_{od}	0.60
d_{v2}, d_{df}, d_{nf}	1.00	d_r, d_{v3}, d_{nf}	1.00
d_{v1}, d_{df}, d_{nf}	0.97	d_r, d_{v2}, d_{nf}	1.00
d_r, d_{od}, d_{df}	0.95	d_{v1}, d_{v2}, d_{v3}	1.00
d_{v1}, d_{v2}, d_{nf}	0.89	d_r, d_{v2}, d_{df}	0.99
d_{v3}, d_{df}, d_{nf}	0.87	d_{v2}, d_{v3}, d_{df}	0.97

Table I.4: *Ngyuen method*, ω

Predictors (Acc)	R^2 median (min, max)	Predictors (AUC)	R^2 median (min, max)
d_{v1}	0.30	d_{v2}	-0.14
d_r	0.28	d_{v1}	-0.18
d_{v3}	0.24	d_{v3}	-0.20
d_{od}	0.23	d_{df}	-0.23
d_{df}	0.20	d_{nf}	-0.24
d_{df}, d_{nf}	0.46	d_r, d_{v1}	0.07
d_{v2}, d_{v3}	0.06	d_{v2}, d_{v3}	-0.26
d_r, d_{od}	0.04	d_{v1}, d_{od}	-0.54
d_{v1}, d_{v2}	-0.00	d_r, d_{v2}	-0.59
d_{v1}, d_{od}	-0.01	d_{v2}, d_{od}	-0.68
d_r, d_{od}, d_{df}	1.00	d_r, d_{v2}, d_{nf}	1.00
d_{v1}, d_{v2}, d_{df}	0.98	d_{v2}, d_{v3}, d_{df}	0.99
d_{v1}, d_{v2}, d_{nf}	0.93	d_r, d_{v3}, d_{nf}	0.98
d_{v2}, d_{df}, d_{nf}	0.64	d_{v2}, d_{v3}, d_{nf}	0.94
d_r, d_{od}, d_{nf}	0.59	d_{v1}, d_{v2}, d_{v3}	0.88

Table I.5: *Ngyuen method*, τ

Predictors (Acc)	R^2 median (min, max)	Predictors (AUC)	R^2 median (min, max)
d_n	0.58	d_n	0.91
d_{df}	-0.11	d_{v1}	0.17
d_{nf}	-0.13	d_r	0.09
d_{od}	-0.18	d_{v3}	0.09
d_r	-0.18	d_{df}	0.09
d_n, d_{v2}	0.77	d_n, d_{v3}	0.87
d_n, d_{v3}	0.69	d_r, d_n	0.86
d_n, d_{v1}	0.57	d_n, d_{v2}	0.86
d_r, d_n	0.48	d_n, d_{od}	0.86
d_n, d_{od}	0.44	d_n, d_{v1}	0.86
d_n, d_{v1}, d_{od}	1.0	d_{v2}, d_{df}, d_{nf}	1.0
d_n, d_{df}, d_{nf}	1.0	d_{v1}, d_{df}, d_{nf}	1.0
d_n, d_{v2}, d_{v3}	1.0	d_{v3}, d_{df}, d_{nf}	0.97
d_r, d_n, d_{v2}	0.95	d_{v1}, d_{v2}, d_{nf}	0.86
d_n, d_{v2}, d_{od}	0.91	d_r, d_n, d_{od}	0.82

Table I.6: Prediction of the method parameters by angular resolution of the databases (M_r) or by the angular resolution and the percentage of vessel pixels in the ground truth (M_{rn}). Least squares method was used to fit the models. *Assessment* reports the segmentation accuracy when the predicted parameters were used.

database	ARIA	CHASEDB1	DRIVE	HRF	STARE
Soares, Λ_{mor}	{2,3,5}	{3,7,9}	{2,3,4}	{5,13,15}	{2,4,5}
M_r	{2,3,5}	{3,7,9}	{2,3,4}	{5,13,15}	{2,4,6}
assessment	93.6	94.6	94.7	95.8	95.1
Azz., s_1, s_2	3.3,1.6	5.3,3.9	2.0,1.6	7.2,6.4	2.8,1.6
M_{rn}	3.2,1.4	4.7,3.6	2.0,1.6	7.4,6.5	3.4,2.0
Azz., r_1, r_2	19,27	21,22	8,25	23,44	12,28
M_{rn}	17.7,24.9	18.0,31.0	7.8,23.4	23.8,40.5	16.0,26.3
Azz., $s0_1, s0_2$	1,0.5	2.5,0	3,1.5	0.5,0	0.5,0.5
M_{rn}	0.8,0.3	1.3,0.4	3.0,1.4	0.87,0	1.4,0.5
Azz., a_1, a_2	0.4,0.1	0.2,0	0.5,0.1	0.4,0	0.5,0.1
M_{rn}	0.4,0.1	0.4,0	0.5,0.1	0.3,0	0.4,0.1
Azz., τ	0.14	0.16	0.16	0.16	0.16
M_{rn}	0.14	0.15	0.16	0.16	0.15
assessment	94.0	94.0	94.5	95.5	95.3
Bankhead, ξ_r	522	780	150	1029	270
M_r	336	634	284	1084	407
assessment	93.8	94.1	93.9	95.6	95.2
Nguyen, τ	1.05	0.95	0.90	1.05	1.05
M_{rn}	1.05	1.01	0.90	1.03	1.01
Nguyen, W	19	33	17	45	17
M_{rn}	18	29	17	46	21
Nguyen, ω	18	16	4	22	16
M_{rn}	16	16	4	22	16
assessment	93.8	94.1	94.5	95.8	95.4